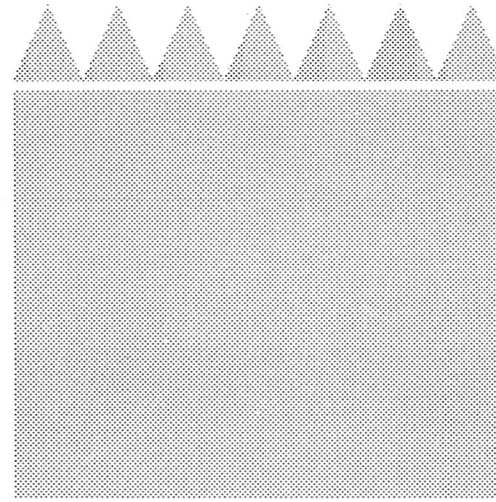
ABS Catalogue No. 1133.0 **NATIONAL STATISTICAL SYSTEMS** A GUIDED TOUR

Australian Bureau of Statistics

NATIONAL STATISTICAL SYSTEMS

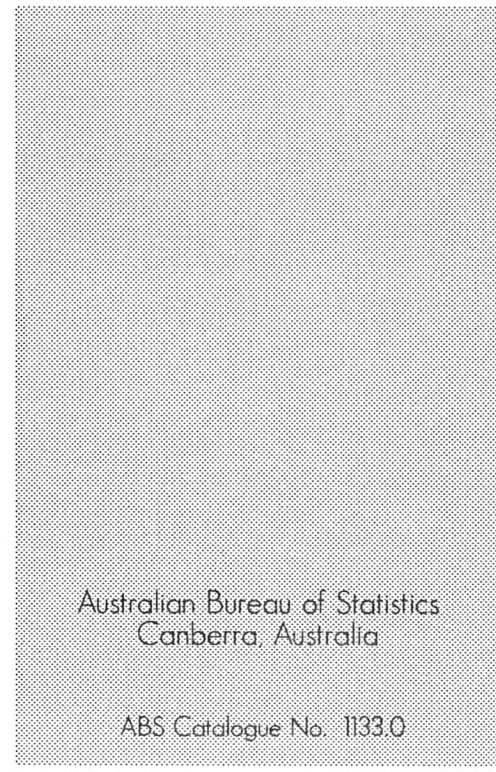
A GUIDED TOUR

ABS Catalogue No. 1133.0



NATIONAL STATISTICAL SYSTEMS

A GUIDED TOUR



Australian Bureau of Statistics
Canberra, Australia

ABS Catalogue No. 1133.0

FOREWORD

"Statistics are too important to be left to the statisticians."

Wassilly Leontieff.

This text discusses the nature, purposes, processes and content of national statistical systems. It was prepared primarily for use in regular courses presented at the Australian Bureau of Statistics (ABS) to staff of ABS and other Australian government agencies but will also be of interest to academic and government institutions in Australia and overseas.

Beyond its obvious relevance to the work of those involved directly in or with government statistical services, this publication is of potential interest to all who have a concern to be better informed about the nature and state of society. The development of national statistical information systems is a collective responsibility and needs to be taken up by the community at large, not just by the statisticians.

The principal author is Ron Fergie, a former ABS officer, who has set out to deal with the topic of national statistical systems in a general and comparative way from a broad perspective of experience with national and international statistical agencies. He was assisted by a number of current senior ABS statisticians and the work was oversighted by a steering committee.

I believe the publication will help to fill a significant gap in the literature.

IAN CASTLES
Australian Statistician

CONTENTS

(A more detailed table of contents is provided in the title page for each PART)

PREFACE

In a nutshell, what are we discussing?

The possibilities of developing national statistical systems in the service of society arise in a great many contexts today. Yet the concept of national statistical systems (NSS) remains ill-defined.

Hopefully this discussion will raise the general level of understanding and interest in national statistical systems as a field of study.

The topic is treated comprehensively here, at least in terms of institutional arrangements, subject matter and processes. The prime intention is to present the subject in overall perspective for those designing, managing and interacting with national statistical systems. It will stop short of discussing such tools of trade as the mathematical techniques of statistics and the technical logic of computing. These are well covered in the literature and fall into place when the nature of the total activity is understood.

Used as a textbook, when supported by case studies, tutorials and other material, the book should provide a framework to contain and relate more specialised statistical training and experience.

Why courses in national statistical systems?

The decision taken by the Australian Bureau of Statistics (ABS) in 1977 to develop the courses which this text supports arose out of a realisation that graduates entering the national statistical service from universities, with majors in subjects like economics or statistics, did not have an adequate intellectual framework or preparation for the kind of work they were to do, and found difficulty in obtaining such a framework on the job. Similarly the role of government statistical services in planning and management was not well understood generally and a need was felt to encompass the topic more systematically than purely internal ABS training had done in the past.

This need was recognised as extending beyond the ABS and the Australian scene insofar as the ABS is an element of a wider system of national and international statistics with which producers and users of statistics have difficulty in coming to grips. Both producers and users of statistics felt a need for a disciplined approach to understanding the nature, objectives, processes and products of national statistical services.

Basic questions such as these needed to be answered —

- How do national statistical information systems work?
- How do users of statistics "interface" with these systems for purposes of solving problems or of influencing statistical programme objectives and design?
- What is the logic of the classifications and definitional frameworks?
- How are the various programmes related ?
- How reliable are the statistical series?
- What do their concepts really mean?
- How do they fit theories about economic and social behaviour?
- How can the users gain access to the detail on the data bases?
- Why are government statisticians so defensive about access to unit records?
- How do they establish their programme priorities?
- What governs the statistical standards adopted around the world ?
- How can the system deliver a continuing timely flow of data relevant to current and emerging problems ?
- How do we identify and extract the data we need from the mass of data available?
- What are the frameworks for understanding and applying the data to our problems?

Content and structure

The topics covered in the four parts of the text are listed in the outline. Part 1 answers the question "*what is an NSS*?" and identifies its basic elements. Part 2 looks at its processes and then its products are examined in Part 3 and Part 4 for the economic analysts and the social and demographic analysts, respectively. The logical development is as follows:

Part 1 — Elements

Here we provide a broad introduction for the other parts, bringing into focus what is special about official statistical services and information systems, in terms of the objectives and institutional arrangements peculiar to government's responsibility for organising mass social data.

It will also discuss the basic elements of such data, so as to provide a working vocabulary for subsequent chapters and to establish the general case for statistical standards.

Part 2 — Processes

This will look at the systems from a process point of view in order to build up a picture of an ongoing developing system. It will start with the process of designing a data collection, collecting data, processing, storing and making available basic statistics in the relatively straightforward case of a large one-off survey. The aim is to bring out, at the outset, basic issues which face statisticians responsible for large scale operations.

The picture of an ongoing statistical system will then begin to take shape as the case studies introduce the further dimensions involved in establishing a sequence of periodic major censuses and more frequent and current indicator surveys — systems derived from a rolling development cycle, linking together current and historical *time series*.

This discussion of integration over time will then be widened and elaborated to consider integration across subject-matter — eg to extend the boundaries of an industrial statistics system to include complementary industries, such as distribution and construction.

This should lead on to an examination of the characteristics of large scale *data base* operations designed to store data in readily accessible form.

Illustrations will then be given of the manner in which this material may be drawn from the data base and brought together and manipulated by analytic processes to produce general-purpose derived statistics such as the national income and product estimates, balance of payments and price estimates.

Finally, the possibilities of identifying and using the body of official statistics to solve problems will be considered and this will be linked back to the question of how the system design can respond to the changes in user requirements that arise from experience in using the available statistical material and from the re-evaluation of the total system and its priorities.

Parts 3 and 4 — Products

Following this study of the system as a process, Parts 3 and 4 will look at the characteristics of the body of statistical information that is produced to meet a nation's goals. This should provide a picture of the major subject-fields insofar as they constitute conceptual systems.

Thus the aim is to link with economic and social theory in establishing standard concepts and classifications relating to the common statistical units (such as households, enterprises, establishments etc), and the data to be collected about them.

The various fields of economic statistics will be discussed in a comparative way in order to see what they have in common and what is peculiar to each. The aim will be to give a comprehensive overview of economic statistics, such as by way of the UN *System of National Accounts*, as an introduction to the standards applying in theory and in practice to specific data collection programme fields.

The fields of demographic and social statistics will be approached differently through a focus on key social concerns and categories of people for which statistical information needs particularly to be brought to bear.

As they examine the broad programme fields and sub-fields in terms of information requirements and standards, Parts 3 and 4 will also consider the nature of the source records and collection vehicles (eg mail, field interview, administrative by-product extraction) applying in each field and influencing the content of the system's output.

Use of the text as a teaching aid

The sequence of parts and chapters is intended to provide a logical structure for studying national statistical systems in the round and then an entree to each subject to be covered.

Individual session leaders may address the same objectives in ways that suit the particular audience's level of understanding and interest or which draw best from their own experience. They may be expected to introduce current case study material to sharpen discussion, but desirably should not divert the general thrust of this outline by procedural detail which may contribute little to an understanding of general principles, issues and themes.

It is envisaged that there would be tutorial sessions aimed at assisting students to grasp more firmly the ideas introduced in the text and lectures and to build on what has been learned by creative exercises, such as the preparation, presentation and discussion of reports, simulated conference and working groups and the like.

Bibliography

The Australian Bureau of Statistics' library (PO Box 10 Belconnen, ACT 2616 Australia) has prepared a select bibliography on National Statistical Systems for use with this text and will update it from time to time. Copies are available on request.

Acknowledgments

A *National Statistical Systems* textbook was initially published under this title in 1977 by the then Canberra College of Advanced Education (CCAE) — now University of Canberra. It took the form of a 600 page 4 volume course book edited by Ron Fergie, the principal author of the present text and written with the help of a number of ABS expert lecturers. It served as a general reference and guide to structure and content for courses held at CCAE and in the ABS.

In 1987 the present much shorter text was prepared for the ABS by Ron Fergie as consultant, initially as an *in house* interim edition. It has been used since then in ABS courses and made available for other courses in the region on request.

In 1990, the economic statistics part of the ABS interim edition was revised and extended by Ron Fergie as a training manual for the United Nations Statistical Institute for Asia and the Pacific. In turn, elements of this have been embodied in the present text, as has fresh material from ABS training courses, particularly in respect of social statistics.

Ron Fergie wishes to acknowledge the expertise and experience of many ABS and other colleagues which have been embodied in the publication. The text has evolved over some 15 years from material prepared for training courses held in Australia and overseas and much helpful advice has been received and contributions from many sources have been incorporated.

PART 1: ELEMENTS

CHAPTER 1
What is a national statistical system?

What is/are statistics?
An end product?
What *is* 'official statistics'?
What *are* official statistics?
Systems, real systems & information systems
The planning and decision making context of an NSS
To sum up

CHAPTER 2
What drives a national statistical system? Its mission, authority, charter and basic needs

The mission — a reason for being
What are the system requirements?
What is the nature of the requirement for authority?
To set up an agency and define its functions
To provide for a post of national statistician
To provide for the collection of statistics
To ensure the security of information supplied
Machinery matters
Some critical balances
Conclusion

CHAPTER 3
The NSS as a process — its organisational elements

Nature of national statistical service activities
Personnel
Facilities
Organisation
Communication and coordination
Conclusion

CHAPTER 4
The NSS as a body of data — standard statistical units, data items and classifications

Standard statistical components
Basic elements of statistical series requiring standardisation
Data item and statistical unit standards
Principles of classification
Origins and levels of standardisation
The process of determining international standards
Establishing standards in practice
In *principle* and in *practice*
Comprehensive standards frameworks
To sum up

CHAPTER 1

WHAT IS A NATIONAL STATISTICAL SYSTEM?

This chapter aims to make clear what we really mean by the term 'national statistical system'. Is there any such thing?

Some would regard an NSS as a sort of mythical being — something we like to think exists but which is really an ideal rather than a fact.

But, whatever the state of a country's statistical arrangements, **the concept of a national statistical system is important to grasp,** both for those responsible for developing a country's statistical services and for those wanting to make use of the services.

Certainly, statistical services need to be purpose-designed systems and users need to understand how best to use the systems and to guide their development.

We will approach the subject by considering what we mean by *statistics* and *systems* and then offer a definition which will open up the subject generally.

What is/are statistics?

One could say that *statistics is the process, art or discipline, of collecting, organising and interpreting numerical data about observable phenomena* — or more simply it is *figuring.*

Some tend to overlook the work of collecting and organising the data and stress the process of drawing inferences from the data — eg *statistics is the process of drawing inferences from numerical data in conditions of uncertainty.*

Such a definition may help to distinguish the work of a statistician from that of an accountant (although it would be a very tidy accounting system that produced all the answers without some need to draw inferences from inadequate figures).

PRODUCTS	*/ PROCESSES*
plans	*planning*
statistics	*statisticulating*

In any event, in the above use of the word, *statistics* **is** a **process**. But if you ask *what **are** statistics*? you will call to mind the ***product*** of the process — *statistics* **are** *systematically organised numerical data*, such as the tables in a statistical publication.

An end product?

but where does it really end?

Many of us are sometimes guilty of thinking of statistics as the **end** product of the statistical process. But statistics, such as tables of figures, are not the end product. If there **is** an end-product it is the **inference** drawn from the statistics, assisting decision-making, research and discussion.

As we study the process of producing and using statistics in Part 2 it should become clear that statistical data are best thought of as statistical material which, when put together and manipulated in various ways, enables more statistics to be produced for the purpose of drawing inferences. Discussion of statistical file systems will show that statistical data are not used up in the statistical process — they can be used and reused to produce new information.

What *is* 'official statistics'?

Here we are talking about that part of the process of producing and using statistics which depends on the actions of governments. When it comes to drawing inferences about society at large it generally requires the intervention of government to collect and organise the data from which to make reliable generalisations.

Opinion surveys and the like may be handled successfully by private initiative, but the marshalling of large masses of data relating to complex fields like prices, incomes, foreign exchange, health or security is necessarily a social act, dependent on the authority and capacity of government.

The word 'statistics' is derived from the word 'state' and much of the original impetus for the development of the discipline arose from the effort to collect data on and draw inferences about nation states. 'The data were not at first numerical... but the precision and convenience of data expressed in numbers have led to... the common use of the term 'statistics' as if it related exclusively to data expressed in numerical form...' (Encyclopaedia Britannica, 1960's edition)

State-istics

The history of statistics as a process has its beginnings in government administrative records which produced the numerical data from which inferences about the nation — its population, taxable capacity, potential for military call-up and so on — could be drawn.

The ancient Egyptians prepared lists of heads of all families and, in ancient Judea, censuses of population were held on a number of occasions. The Romans held 70 censuses including the one we read of in the Scriptures.

We will focus on what it is that characterises the activities of government in the process of producing and using statistics. We will analyse the activities of the government statistical service as a specialised activity within the overall national planning and decision-making process.

Clearly goverment statistics as a process is a service function, providing information about the nation which only governments will have the power and capability to provide efficiently. It is a service to the community as a whole, not just a tool of government.

the right to know

Access to statistical information is increasingly regarded as a fundamental right of society. The government in power needs this to perform its functions of economic and social planning and management. The Opposition cannot exercise its critical role without it.

Private enterprises, and the whole international community necessarily depend on the availability of official statistics as a framework for their plans and actions.

What *are* official statistics?

We have said that statistics are systematically organised numerical data. Official statistics are those provided by governments. The word *official* may suggest that the data are authoritative in the sense of being absolute truth and are in some way certified *facts* once they are released by the national statistician.

facts? or probabilities?

Responsible Government statisticians would have to point out that they deal generally with *probabilities* rather than facts and that, like other statisticians, they are in the business of trying to draw inferences in conditions of uncertainty.

They necessarily use judgement in producing official statistics and considerable risks of error are involved at every stage through the collection and processing of data to produce basic statistical tables. As Walter M. Perkins of the US Bureau of Statistics put it, *"it should be apparent that it is the sheerest coincidence if a published statistic happens to correspond with the unknown truth."*

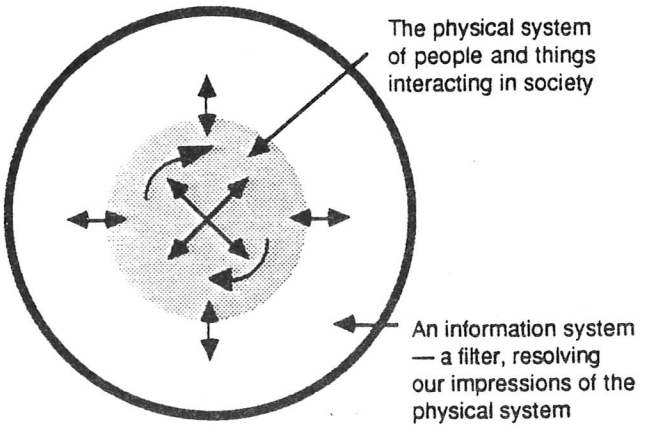
The core of the problem is that official statistics, more than other statistics, can generally only be a representative description of the physical events behind them. Statistics covering large communities are rarely obtained by direct observation; they are made up of information obtained second-hand from the people and businesses who supply information from recollections or records they keep about their activities.

Systems, real systems and information systems

A system can be defined as a group of interacting, interrelated, or interdependent elements forming or regarded as forming a collective entity.

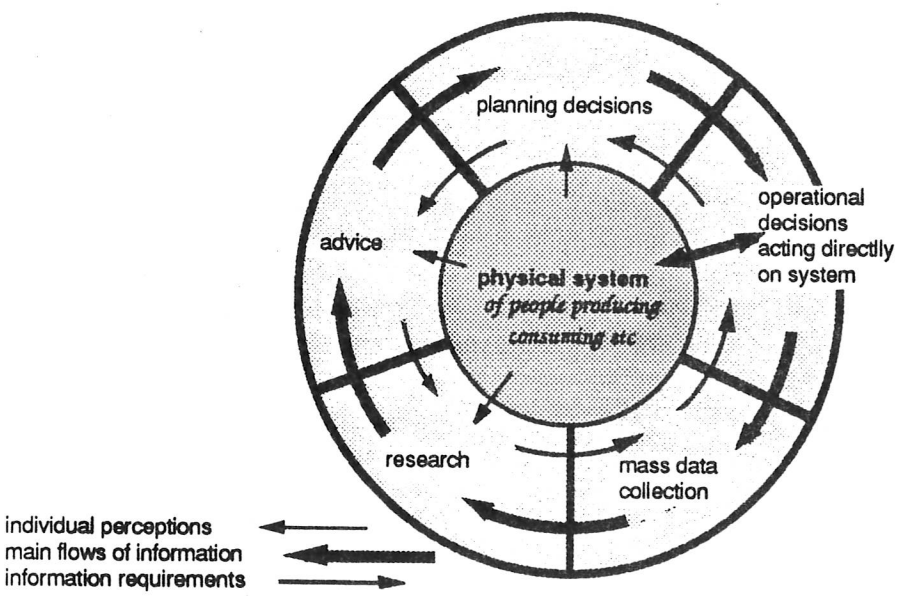
It is useful to distinguish then between a *'real'* physical system and an information system **about** that system.

In a *real system* people in a community are interacting in buying, selling, building etc. Associated with this is an *information system* through which we gather and resolve data about these physical activities. This is done by a variety of means, including perception by individuals, the organised censuses and surveys, record-keeping of government and the analysis of the data by various institutions and specialists.



The more complex the social group the more necessary it is to rely on indirect large-scale data collection and processing operations rather than on individual perceptions.

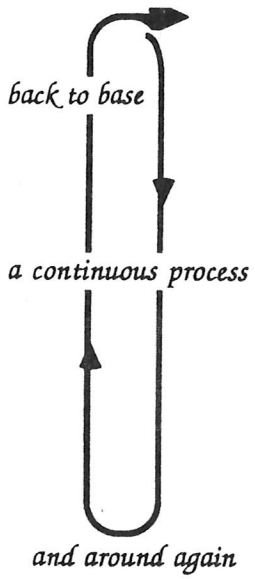
Even the most talented observer can hardly make reliable generalisations about (say) the income or employment of a large community, except with the help of the kind of social arrangements for marshalling information illustrated in the following diagram.



This pictures the very loose, unstructured kind of system within which a government statistical service operates. It shows the flows of information between the various broad functional specialisations involved in sensing what is going on in the community, and in reacting to it.

The planning and decision making context of an NSS

In the above model, the function of mass data collection performed by a government statistical service is shown as an important link in a continuous circular sequence as follows:



operational decisions (by the community to buy, to sell, to build, to migrate, etc)

mass data collection (by an official statistical service observing these activities, recording and summarising them — the processes of a national statistical system)

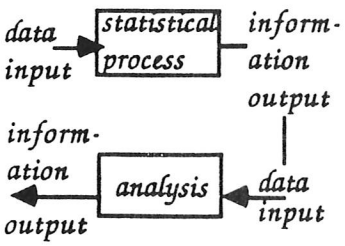
research (eg by academic institutions digesting statistics in forming theories about the causes of and connections between these activities)

advice (eg by economists and others using the theories and the current statistics)

planning decisions (eg by governments and key executives who consider the advice and determine the policy and programmes which establish the framework for operational decisions)...and so on.

Operational decisions are made day by day within this framework and (eg when expressed as transactions in the records of enterprises and governments) form part of the basis for the next round of data collection, analysis, interpretation, advice, policy decision — as illustrated.

There would also be subject-matter divisions superimposed on the functional divisions of the 'system' (eg *economic* research, *social* research, *environmental* research) and one could imagine all sorts of cross-flows of information (eg once theory and models are established much of the statistical data will flow direct to the decision-makers and their advisers rather than by way of the research function).



Note also that data become information only when they are in a form (often as a result of processing) which meets the information requirements of the person receiving them. Thus the statistical tables of the Government statistical publication may be information for researchers but may not be information for policy makers until they are further processed into the advice they need to solve the problem at hand.

In the circular flow of information pictured, inputs of data are being converted into information so that the output of one processor will become the input of another and there is a feedback of requirements, after the manner of an automatic control system.

So it will be a basic theme of this 'guided tour' of national statistical systems that producers and users of statistics, specialists working at all levels, should constantly be looking over each other's shoulders. We need good communications across the specialised functions in the statistical process, across subject-matter fields and between the geographical and political divisions of the international statistical system (e.g across local, State, national and international institutions).

Thus the term *national statistical system* is intended here to refer to that part of the overall information services available to a nation which is the specialisation of government statistical services.

It refers to both the systematically organised *body of data* and to the system of *mass data collection and processing* which creates and maintains it and makes it available. We will aim to put this product and these processes into proper context within still broader planning and decision making systems.

It will be argued that the task of developing better information systems to solve economic and social problems is not the Government Statistician's alone, but is a collective one. Reference to a country's arrangements for doing this as a 'national statistical system' at least signals an intention to ensure that statistical programs really 'hang together' as instruments for pursuing community goals.

To sum up

It is worth underlining four major points —

- ***socio-economic planning and management of any complexity can hardly function without a statistical system to support it***

- ***the scale of this function of mass data collection will involve the authority and resources of government***

- ***the ultimate objective of government statistics is to help solve problems*** *— not just to produce statisics*

- ***the functions depicted in the diagram at page 5 are all interdependent and require constant two-way communication*** *— in the case of an urban community, for example, the objective of each of these functions must be to contribute to the process of rational decision-making affecting the quality of life in the society. If the mutual interdependence is not clearly recognised in the interactions between the components of the system, information will not flow in response to need and the planning and management of cities will tend to break down.*

CHAPTER 2

WHAT DRIVES A NATIONAL STATISTICAL SYSTEM? ITS MISSION, AUTHORITY, CHARTER AND BASIC NEEDS

Statistical systems of national scope may grow 'like Topsy' out of the efforts of various government agencies and the efforts of (say) a National Planning Office to pull these together in successive National Development Plans or other periodical appraisals of the state of the nation.

More often the drive comes from a central statistical office which is given the legislative authority and charter to establish or co-ordinate a national statistical system and does so in accord with the way successive heads of the central statistical organisation perceive the task.

The authority and charter of major statistical agencies or of agencies which co-ordinate government statistics are likely to be set out in statistics legislation. The manner in which the role is to be performed may emerge as a departmental philosophy documented progressively in policy statements and manuals, annual reports and perhaps a corporate plan setting out longer term goals, objectives and strategies.

The mission — a reason for being

A good statistical system will certainly be one developed with clear purposes in mind. Statistical legislation commonly spells out the role of a government statistical service, by listing the kinds of statistics which it is required to produce.

intermediate objectives

Some countries have found this an unsatisfactory approach to expressing the mission of a national statistical organisation in that it fails to bring out the fact that data collections are really only intermediate objectives — the dissemination of statistics being only a step in meeting more ultimate objectives of policy making, plan formulation and management in accordance with national goals.

ultimate purposes

An alternative is to specify the organisation's mission in the very broadest terms before going on to spell out the subjects it should cover or the particular functions it should perform.

An example of such a broad statement of mission is 'to assist and encourage informed decision-making, research and discussion within governments and the community, by providing a high-quality user-oriented and dynamic statistical service; we will actively co-ordinate statistical activities across government agencies and promote the use of statistical standards' (Australian Bureau of Statistics, Corporate Plan, 1991).

What are the system requirements?

However extensive and demanding the Government Statistician's mission, some of the basic things that are necessary for a government statistical service to function properly will be:

- *authority*
  - of law
  - of respect

- *public confidence*
  - in the worth of what it is doing
  - in its respect for the confidentiality and privacy of the suppliers of data
  - in its professional competence
  - in its impartiality and objectivity

- *personnel* — well trained, dedicated and organised for efficient operations

- *facilities* — for large-scale data processing operations

- *standards*— to ensure the integrity of its product in terms of concepts, definitions, classifications and operational methods

- *philosophy* — as to the nature of its service role.

What is the nature of the requirement for authority?

The real authority of a statistical agency lies in the goodwill, competence and tact through which it obtains the respect and cooperation of suppliers of information. But legal authority is necessary for any government agency to function and the power to require people to supply statistics as a public duty generally needs to be legislated for.

Thus a clear delineation of a statistical agency's functions, powers and responsibilities is a valuable framework for its effort to enlist the cooperation of users, respondents and other government agencies. Some of the basic elements that one should look for in statistics legislation are discussed below, together with some of the important issues that arise.

To set up an agency and define its functions

The legislation should identify the agency and indicate its functions. These functions will differ according to whether it is for a central statistical office which operates

permanently as a coordinating and standard-setting agency, (eg the Central Statistical Office in the UK); as a centralised collecting agency (eg the Bureau of the Census in the US) or as a combination of the two as in the Australian Bureau of Statistics (ABS).

Thus, in the case of Australia, powers, functions, duties and responsibilities of the central statistical agency have been expressed as follows in the Australian Bureau of Statistics Act, 1975:

(a) to constitute the central statistical authority for the Australian Government and, by arrangements with the Governments of the States, provide statistical services for those Governments;

(b) to collect, compile, analyse and disseminate statistics and related information;

(c) to ensure coordination of the operations of official bodies in the collection, compilation and dissemination of statistics and related information, with particular regard to:

(i) the avoidance of duplication in the collection by official bodies of information for statistical purposes;

(ii) the attainment of compatibility between, and the integration of, statistics compiled by official bodies; and

(iii) the maximum possible utilization, for statistical purposes, of information, and means of collection of information, available to official bodies;

(d) to formulate, and ensure compliance with, standards for the carrying out by official bodies of operations for statistical purposes;

(e) to provide advice and assistance to official bodies in relation to statistics; and

(f) to provide liaison between Australia, on the one hand, and other countries and international organizations, on the other hand, in relation to statistical matters.

Some of the issues arising include:

- is the coordination role to emphasise the negative side of ensuring that other agencies take no initiatives without consultation? — or the positive side of providing advice and assistance, minimising duplication, setting standards and promoting and facilitating development generally?

- will the aim of ensuring that the total stock of data held by governments is fully used run into difficulties because of legislation controlling access to data in other agencies? (eg under a Taxation Act the taxation returns held by a Commissioner of Taxation may not be available to other agencies of the government)

To provide for a post of national statistician

This section will provide for the appointment of a National Statistician, indicating duties, powers and tenure in the position. Questions which arise here include:

- should National Statisticians be required to have professional qualifications?

- how secure should their tenure be?
  eg dismissed only by the Head of State on specified grounds?

- how independent should they be?
  - attached to a Department and answerable to the Departmental Head?
  - answerable direct to the Minister? — solely responsible for decisions on procedures, methods and the extent, form and timing of publication of statistics?
  - answerable to Parliament by way of an annual report (with an opportunity to report any interference with their professional independence or any serious failure by other agencies to cooperate in developing efficient and cohesive statistical systems)?
  - able to recruit their own temporary staff?

To provide for the collection of statistics

The legislation may need to define what is meant by *statistics* and the ways in which it is proper to collect them. Thus:

scope of NSS

- should there be a list of topics in respect of which the Statistician is required to collect and publish statistics?

- should there be a requirement to undertake (eg) a population census at specified intervals?

- is it necessary to state that statistics can be collected by sampling methods?

- is it necessary to state that statistics may include opinions, such as estimates of expected capital expenditure?

- should there be specific provision for data obtained by other agencies for non-statistical purposes to be passed on to the statistical agency?

- should there be joint arrangements with other agencies for collecting statistics?

- should the Act apply to all surveys conducted for statistical purposes by agencies of the State (ie with its general provisions as to confidentiality applying to the surveys and with the Statistician responsible for approving them?)

privacy issues

- should departments, businesses and other respondents be required to give the Statistician's officers access to their premises and records?

- what are appropriate penalties for non-compliance by respondents or for offences by employees of the statistical service?

- what should be the relation between statistics legislation and that relating to other information topics — eg on *freedom of information, privacy, archives, business regulation, paper burden.*

To ensure the security of information supplied

This section must give quite clear and positive assurance that information collected under the statistics legislation shall be used for statistical purposes only.

how confidential?

A basic issue that will arise is whether there are any exceptional situations in which information relating to individual respondents may be released to persons who are not employed under the legislation and are not bound by its secrecy provisions. For example, to provide for:

- release of information which is available from other sources

- the release of information for which consent to release has been obtained

- the release of information which cannot be related to any identifiable individual (eg when identifying characteristics have been deleted)

- the release of valuable but generally non-sensitive information, such as an index or list of businesses with minimal information as to the nature of their activities and their size in terms of employees

- the possible exceptional case where it would be clearly in the national interest to publish statistics identifiable as those of an individual business (eg where one or two businesses dominate an important

industry and it would not be possible to publish the figures for that industry without revealing the figures supplied by an individual business)

- the concept that 'disclosure' provisions should allow sharing of data among all persons subject to the statistics legislation

- the concept of deeming persons employed to perform special services under contract etc (including consultancy advice, the conduct of research or provision of data processing services) as being subject to the legislation's security provisions.

Machinery matters

It is important that the intention of the legislation should come through clearly and not be confused by detailed drafting to cover legal loopholes etc. It is generally the practice to cover these matters in a section on definitions of terms, penalties, delegations, provision for the elaboration of the Act through the making of regulations consistent with the Act, and various other sections covering matters of procedure which are important for the smooth administration of the Act.

Some critical balances

The above discussion of the elements of statistical legislation raises some crucial issues of departmental philosophy which are touched on below and will concern us at various points in this *guided tour*

Legalism or Realism

The drafting of legislation will reflect current circumstances such as needs for information, technical capabilities and the prevailing climate of opinion on the rights of the individual (eg to privacy) versus the corporate rights of the community to be informed (eg as to the activities of other individuals and the government).

But many statistical agencies have to live with out-of-date and patched-up legislation. In that case, does one follow it strictly in both letter and spirit or 'bend it' where it seems reasonable to do so? Statistical agencies which have taken an easy-going attitude to such matters have found themselves forced to bend under pressure and their credibility has suffered.

Responsibility versus professional Independence

Clearly the Statistician is a public servant and subject to the principle of Ministerial responsibility or its equivalent. At the same time, as a policy-neutral agency within the bureaucracy, it is vital also to be in fact and appearance, objective, impartial and professionally independent.

Government statisticians inevitably walk a tight-rope in preserving this balance, with or without well-drafted statistics legislation.

The Need to Know versus private and public Costs of Supplying Information

Statisticians seek to balance the interests of suppliers and users of data using their technical capability to the full — in the long run their systems are dependent on the confidence and cooperation of the community as both suppliers and users of the information which they process on the community's behalf.

The community of users may argue a desperate '*need to know*'. But suppliers of information will need to be convinced that their rights to privacy and confidentiality are fully respected and that demands made on them are reasonable, in the light of the uses being made of the statistics. As the community's perception of these rights and the appropriate balance between them will differ from country to country and over time, the lines are never easy to draw in practice.

Conclusion

The following quotations from United Nations documents effectively sum up this discussion of the objectives and philosophy underlying national statistical services —

"The demand for statistics of the highest quality by Governments, business, economists, social workers and others, has increased enormously. This is accounted for partly by the growing complexity of the modern world, but even more by the fact that many Governments today have accepted wide responsibility for the welfare of the citizens and have embarked upon a wide variety of (economic planning, regulation and social security) measures. To function efficiently, such Governments must have a basis of sound statistical information to assist them in formulating their policies. Governments which do not have well developed statistical organisations are severely handicapped"

"In all fields, both national and international, the work of the statistician is fundamental for comprehending the numerical aspects of the problems to be dealt with, for setting more or less isolated phenomena in their proper perspective, for indicating the significance of parts in relation to the whole and for substituting realistic data for wishful thinking, biased claims and political oratory. In short, statisticians help to shape economic and business policy, they furnish navigational guides for the ship of State, they help social agencies which work for the welfare of the people and they help students and the general public to have a more realistic view of the complex economic and social environment in which they live"

"The statistical service must be objective and impartial in all aspects of its operations, and in particular in the content and release time of its publications. It must be immune to special influences and the results of censuses and surveys should be published at the earliest possible time. Like the judiciary of a country, it must stand above any special interest groups. Moreover, it must be seen to be that way. Otherwise, its external capability and indeed its utility are undermined. In the maelstrom of changing economic and social events and views, it is a towering national asset to have a service which displays the facts regularly in an objective and orderly manner."

CHAPTER 3

THE NSS AS A PROCESS — ITS ORGANISATIONAL ELEMENTS

It is one thing to legislate for a central statistical office to establish a national statistical system. It is quite another to assemble the necessary personnel and facilities into an efficient organisation to do the job.

It is not simply a matter of setting up a central statistical office. For example, since there can be a statistical aspect to every field of activity in the machinery of government, the national statistical system necessarily embraces many agencies in addition to the central statistical office. How then do we ensure that all these elements will, in fact, mesh together into something that deserves to be referred to as a 'national statistical system'?

In this chapter we consider the personnel and facilities involved in a national statistical system and some of the different approaches to organising a statistical service.

Nature of national statistical service activities

The responsibilities of a national statistical organisation, as spelled out in the previous chapter, imply:

- **Large scale, general purpose, repetitive and interconnected statistical programmes** — this can involve a large number of staff (eg about 60,000 in India, 3,500 in the Australian Bureau of Statistics, 100 in Papua New Guinea's National Statistical Office) plus thousands of 'extras' from time to time for major censuses;

- **Inter-agency arrangements to promote, coordinate and set standards** — getting the most from the total stock of data collected for administrative purposes and supplementing this by direct mail and field censuses and surveys. This calls for well qualified professional staff who understand the nature and use of the data in different subject fields and the techniques of acquiring and processing information;

- **An enormous mass of data to process** — this calls for a heavy investment in data processing equipment and systems, e.g. the ABS' computer facilities are among the largest in the country, while the justification for much of the computing capacity in other agencies is statistical in nature.

Personnel

In a large modern statistical organisation it is possible to distinguish three broad kinds of activity going on and calling for somewhat different backgrounds in academic qualifications and experience. These activities are:

designers and analysts

- **subject matter research and analysis**. This is the activity in which information requirements are determined, the output of statistical programmes is analysed and statistical information is computed by analytical means. It requires expertise in the subject matter of the statistical systems at least as much as expertise in statistical methods and it calls for close liaison with users of statistics. Most members of the teams involved in this activity will have majored in such disciplines as statistics, economics and demography. The knowledge required here is apparent in Parts 3 and 4.

producers of statistics

- **data collection and processing operations.** In this activity statistical programmes are designed and managed by statisticians who are expert in methods of collecting and processing data. The professionals who move most naturally into this work are those with academic backgrounds and interests centred on such topics as computing, mathematical statistics (including sample design) geography and the like. The knowledge required is indicated in Part 2.

back-up support

- **general support services**. The direct statistical activities must be supported by such services as personnel and financial administration, computing, printing, publications and library, performed by persons with a general background in administration and computer operations. Naturally it helps if such functions are staffed by professionals who also have some background in subject matter research and analysis and in data collection and processing design and operations, but in a large organisation, staff tend to find a permanent place in one or other of these types of work.

Facilities

There are considerable economies of scale in data processing, storage and information dissemination facilities and this tends to encourage a centralised or (at least) highly integrated approach to the organisation of statistical services. What is important then is not so much the physical location of the central computer as its accessibility. In well equipped services, ready access is assured from any part of the network.

economies of scale versus independence

In some smaller countries, computer support services have been centralised for all government purposes, including statistics. This can raise problems in connection with the Statistician's control over timetables and over the security of the data collected, but may become less of a problem with greater reliance on the more powerful desktop computers now available.

An intermediate position may be to concentrate data processing for administrative by-product statistics in a government computing centre, while providing the Statistician with sufficient computing capacity to process his direct censuses and surveys and to undertake analytical work, including preparation of national accounts.

Organisation

In general terms *organisation* deals with various ways to facilitate efficiency (to do things *right*) and effectiveness (to do the right *things*).

Most essentially a statistical organisation should do the right things, ie produce the information most needed in the form required by the users of statistical services. The problem in practice is that for historical and other reasons it is often easier to operate more efficiently doing the wrong things than it is to do the right things. So it is important to develop organisational arrangements that ensure that the statistical system is not only efficient but is also responsive to the changing information requirements of users.

A central problem in organisational arrangements is how to divide up the total task between the staff and facilities available. This involves the establishment of effective communication between the statistical system and other parts of the total national planning and management information system (as visualised in Chapter 1) and effective communication within the system. Broadly speaking, the organisation of a statistical service has three dimensions — by geographical area, by subject matter and by function ie in terms of **where** you do the job, **what** you do and **how** you do it.

Geographical Organisation

In a national statistical system the principal factor here will be the number of levels of government, the extent of their powers and the relationship between them.

In Australia, for example, State Governments had statistical offices before Federation and the integration of these by Commonwealth/State agreements resulted in a substantial geographical centralisation of personnel and facilities in Canberra. However, large operational centres are maintained in the State capitals and service both State government and national statistical programmes.

In New Zealand, on the other hand, there is no federal system of government. National statistics were centralised in the Department of Statistics, Wellington, from the beginning and, while there has since been a decentralisation of a number of statistical fields to Auckland, Christchurch and Dunedin, the overall control is still firmly based in Wellington.

influence of computers

While political considerations may ensure that a large part of the data collection function will remain decentralised, and some part of the research and analysis function will be associated with government at local and provincial levels, data processing tends to be geographically centralised at the points where automatic data processing facilities are established on an increasingly larger scale.

Counterbalancing the centralisation pressures, however, is greater accessibility from all points in the system via desktop computers, subject, of course, to appropriate security controls. At the same time the planning and management of systems tends to be centralised as they become national in scope and common conceptual and operational standards are necessary.

Subject-matter Organisation

subject-matter centralisation

The terms *centralised* or *decentralised* as applied to statistical organisations more often refer to subject-matter centralisation rather than to geographical centralisation.

Thus the Australian national statistical system has always been described as a highly centralised system in the sense that most of the data collection and processing over the full range of economic and social statistics has been undertaken by a single organisation.

decentralisation

If the statistical service were divided up among a number of agencies responsible for different subject-matter (eg commerce, health, education, agriculture, labour and so on) and the statistics relevant to these, we would have a decentralised system which is held together in terms of common standards etc by a relatively small central statistical office. Most of the larger countries, such as USA, USSR, Germany, UK, Japan, Philippines and India, are more or less decentralised.

On the other hand Australia, Canada, New Zealand, Netherlands, Sweden and many other intermediate sized countries and most of the smaller developing countries have a high degree of centralisation, with a single large national statistical office, which also exercises a coordinating role over such relatively small statistical organisations as are established in other agencies.

centralisation = integration?

Advocates of the centralised system point to the desirability of promoting the conceptual and operational integration of the system. Generally this is easier when the statistical service is under one management (but since the large task still has to be divided up within that organisation, there can still be a lack of integration due to poor internal communications and management).

Arguments for the decentralised system stress that the initiative for producing statistics should be as close as possible to the interested users.

Not being users themselves, large monolithic statistical organisations are likely to be more efficient, but less effective in producing what is needed, when it is needed. **However, new electronic facilities for linking and accessing data bases are tending to support both system integration and decentralised user services.**

The case for centralisation seems to be strongest in respect of economic statistics, where the body of theory governing the use of the data is relatively unified and a parallel general-purpose data-organising system is well established in the form of the UN System of National Accounts. It is also strongest in the case of small developing countries where the statistical units of most agencies would generally not be large enough to have the 'critical mass' of facilities and stable professionally qualified staff.

'leavening the lump'

The difficulty of keeping professional staff in small units can be met in a decentralised system by having a class of professional statisticians in the public service who can move from agency to agency but remain in statistical service functions. This is the approach adopted in the UK.

The problem of providing the decentralised statistical offices with adequate data processing facilities is not easily met in a small country. Even countries, such as the UK, which hold fast to the principle that subject-matter statisticians should continue to be located in the Departments, may centralise production of economic statistics around central registers of employers and unified data storage facilities.

Within the statistical office in a centralised system, such as in Australia, subject-matter specialisation can still follow the Government's division into ministries or Departments. This categorisation by the nature of the information requirement (eg about industry, health, education etc) needs to be done with care to ensure that the logic of the division is clear and consistent and does not become out-of-date with every shift in national policy emphasis.

Functional Specialisation

The third dimension in statistical system organisation has to do with the different kinds of activities to be performed.

We have already considered (under *personnel*) division by task specialisation which would split the work of a statistical organisation into (say) a *research and analysis division*, a *system design and operation division* and *a general support services division.* Some national systems will concentrate the last two functions into a large central statistical agency and leave the Departments to concentrate on research and analysis.

Fully decentralised statistical systems have all three functions located in Departments but react to the new technical developments by concentrating the large scale general-purpose data collection and processing programmes (such as for the population and economic censuses) in a major statistical Bureau which is sometimes also the central standards authority.

task specialisation
— efficient
— effective?

The main case for dividing up the statistical organisation by function rather than by subject-matter is that it enables distinct types of tasks like field survey design and operations and computer systems design and programming to be undertaken on a larger scale, with economies in specialised skills and facilities.

The weakness is generally that objectives are not as clearly seen as in a subject-matter division of operations which helps to ensure that the demand for information in specific fields is clearly understood and appropriately met. With responsibility for information in each subject field clearly identified, lapses in data quality are less likely to happen.

In practice there is likely to be a mixture of subject-matter and functional organisational forms with the objectives of obtaining the advantages of efficiency possible from functional division of the task while having a subject-matter specialisation that is sufficient to ensure effectiveness.

who is responsible for what?

It is of some importance that, whatever the complexity of a statistical agency's internal organisation arrangements in collecting and processing statistics, the responsibility for specifying information requirements and analysing output should be readily apparent to users.

If, for example, much of the economic data (as well as social data) in respect of the household sector is collected under different collection programmes, such as a population census branch, a field survey branch, a finance and taxation branch etc, there should also be a clearly identifiable subject-matter unit which has a comprehensive grasp of the various sources of data in respect of the economic activities of households. Equally it would be desirable to be able to identify subject-matter responsibility for a basic kind of activity, such as production, whether this be by enterprises, households or government and non-profit agencies.

One device that may help overcome this kind of problem is to produce a comprehensive publication, such as a Year Book or a Statistical Register in which the statistics from the various programmes are associated under appropriate subject-matter headings arranged in a pattern that best reflects the way users would associate them.

Communication and coordination

In a statistical organisation of any size there will be a natural tendency for specialist units to emerge on mixed subject-matter and functional lines.

All the technical advantages of having e.g., a centralised field survey organisation and a centralised data processing organisation skilled in computer systems for processing and storing the data, lead to the situation in which no division or branch can carry out its statistical programmes unilaterally.

If communications and coordination are not good, if there is not a general appreciation of the organisation's ultimate objectives and a common mind about how to achieve the objectives, the organisation will not be effective, however logically organised it may be.

As a classic report to the UN Statistical Commission puts it:
'In particular, subject-matter personnel may experience a sense of loss of control over the outputs. As methodologies and systems become complex and highly specialised, those familiar with them assume strategic positions in the statistical production process. When the functional specialists use their knowledge unwisely or thoughtlessly, serious problems arise. Coupled with this is the danger that the systems and methodologies may assume a life of their own with insufficient attention being given to the needs of the subject-matter outputs for which the methodologies and systems are merely a means. It is worth noting also, particularly in regard to computer use, that the subject-matter experts may not be sufficiently aware of the vital importance of clearly specifying their output requirements in detail. As a consequence, the computer specialists may make decisions on the basis of incomplete or faulty information resulting in inconsistent or unusable outputs.'

UNITED NATIONS. Statistical Commission, 19th New Delhi, 1976, "Statistical Organisation: the organisation of statistical services; review of major issues"; report by the Secretary General. UN Document E/CN/3/495 para 27

Since the 1970's subject -matter personnel have acquired more direct control through technological changes which have put personal computers at their desks and linked them much more clearly with the *automatic data processing* environment.

communication again!

But there remains a need to ensure that there is good communication across the lines of a statistical office's formal organisation and that a good general background is established by training in its various forms. There are many devices for doing this through the formation of inter-branch project teams for major new programmes and through ad hoc or regular working groups. All of this requires strong direction from the top management of the statistical agency.

The problem is essentially the same whether the statistical system is centralised or decentralised, except that in the latter case the devices for coordination and communication may have to be more formal.

Thus some decentralised systems have a central office of statistical coordination which concentrates wholly on this task and has no statistical data collection and processing operations under its roof.

Most statistical services are likely also to have formal statistical advisory committees to improve the link between the statistical system and users. One would expect that this would be particularly desirable for large centralised statistical agencies which are not themselves users of statistics.

Conclusion

This chapter has touched on only a few of the many aspects of the problem of organising statistical systems for the purpose of optimising efficiency and effectiveness. It has demonstrated:

(a) that the work of a statistical service involves a range of expertise beyond the 'statistical' skills that are likely to be the prime concern of the teaching of statistics in most academic institutions.

(b) that formal organisations for employing these skills are only frameworks and are not likely to work effectively without good communications across organisational lines. This 'guided tour' will continue to stress the paramount need for good communications. Issues arising in ensuring good communications, within a government statistical service and between this and the other elements of the total information system of which it is part, will come up again in subsequent discussion.

CHAPTER 4

The NSS as a body of data— standard statistical units, data items and classifications

We have so far discussed some key elements that make up a national statistical system, viewed as machinery for producing statistical information. In this chapter we focus on economic statistics as a product and ask 'what constitutes a systematically organised body of data?'

The discussion will be general and introductory only, since issues involved in integrating statistical programs will be further discussed as they arise in covering the statistical development process in Part 2, while the particular statistical standards applying to different subject-matter fields will be discussed in Parts 3 and 4.

Standard statistical components

The idea of using standard components is understood in respect of physical systems like a building. It is not as immediately obvious how it applies to an information system and what is meant in this context by terms like *comparability, integration, consistency, interrelatability, coherence* and other similar buzz words which seem to refer to highly desired attributes of statistical series making up a statistical system.

Users of statistics soon get the idea, the hard way, as they find it necessary to use a number of different statistical series — about labour, prices, production, imports, exports and so on — in association with one another, in order to pursue a particular line of enquiry or to produce new statistical information which is more directly relevant to the particular problem being considered.

will the figures 'hang together'?

As they dig for data, the users hope fervently that the available series will, in fact, *hang together* and that they are working with a system of statistics which can be manipulated confidently, with some assurance that the definitions, classifications and reliability of measurement are what the users might reasonably expect them to be.

So users need to know the specifications of the statistics they propose to use. What then are the elements of a statistical series in respect of which users will need specifications?

Basic elements requiring standardisation

'Things' and their quantitative values

There are of course, two basic elements to every statistic: there is the description of the thing being measured and there is the quantitative value to be attached to it — the figures.

Mathematical statistics provides the technique for analysing the figures. But in this chapter we are concerned about the literary (as opposed to quantitative) side of statistics; the side that concerns concepts, the things being measured and the way in which they are described.

what things?

Don't underestimate the importance of this matter of language, of ideas and their communication. No amount of precision in figuring can compensate for lack of clarity in defining the item for which data have been collected.

And it is not just a matter of devising beautifully elegant descriptions of a concept which is to be the subject of measurement. We might devise a completely satisfying and logically defensible definition for *shipments of goods* (as an example) but it is not much good if the businesses reporting this information mis-understand the definition and give figures for something else, which they think is *shipment of goods.*

whose perception of the thing?

It is a core problem for statistics that there can be very severe problems of communication between the user of statistics, the producer of statistics and the people who supply the original data. The words describing a statistical item on a questionnaire and in a publication may mean different things to each user, producer and supplier. And the resulting statistics may not be what they appear to be.

3 dimensions of a thing

There are many possible dimensions to the description of a data item being measured in a statistical collection. It is customary to think of the data (each item in a table) as having three standard dimensions. Data must be represented by a reference to the object to which the facts are connected, descriptions of the facts in question and time specifications.

1. the object or statistical unit

The object or set of objects to which the facts refer is called a *unit* or *statistical unit.* The statistical unit may be a person, an establishment, an enterprise, a country — whatever types of unit it is the practice to collect statistics about.

2. the data item

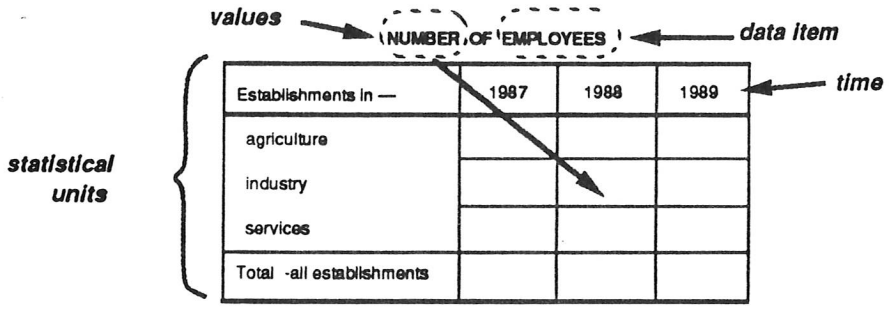
The description of the fact may be divided into the characteristics or *data item,* which describe the kind of fact recorded and the *values*, of the figure which has been observed or computed in respect of it.

3. time specification

The last dimension is the *time* specification.

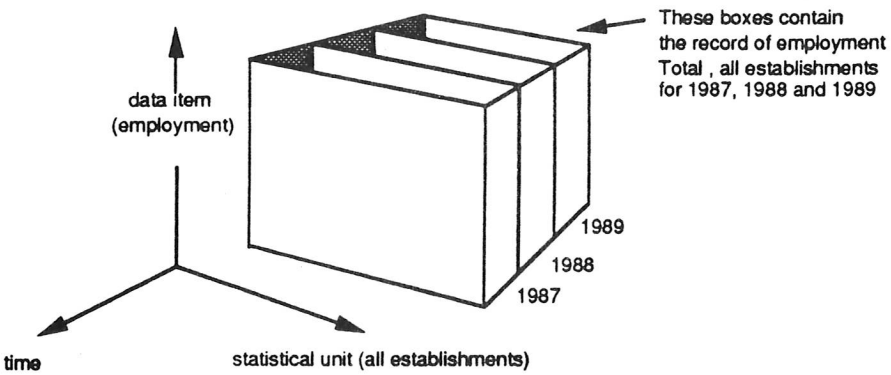
Thus in a typical statistical table, the facts relating to different kinds of statistical units (eg establishments in agriculture, secondary industry etc) in different time

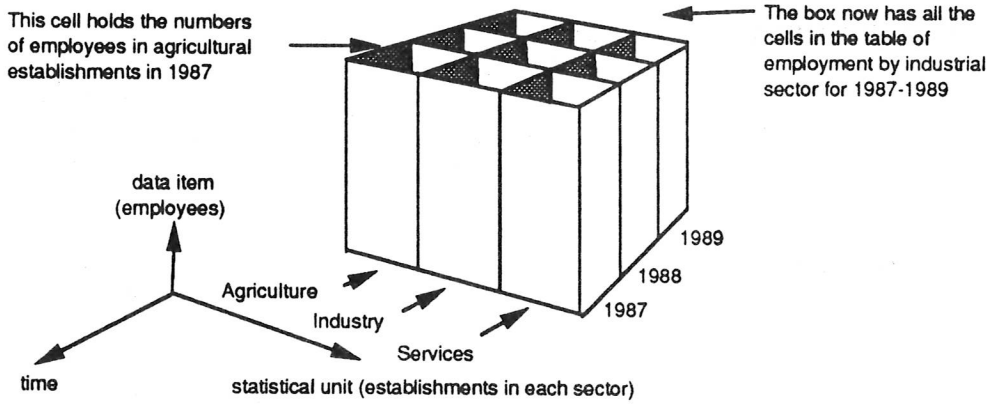
periods are described in terms of a characteristic or data item to which numerical values are attached (eg the number of employees of these establishments).



Such a table illustrates the objective of reducing raw collected data into totals for common categories or *classes* in order to provide information in the form and level of detail that users can assimilate.

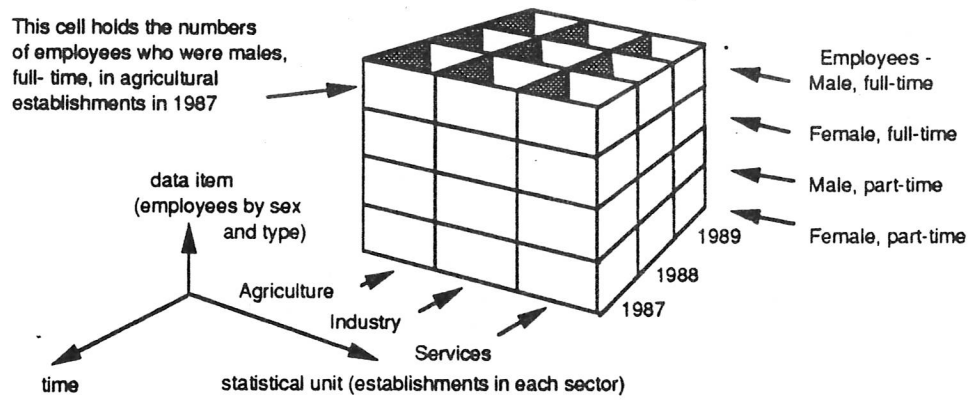
It is useful to think of the cells in the table (or in a computer memory) as a *data box* containing a number of rooms for showing values. What we are doing then is to sort the total survey *population* of statistical unit records into appropriate classification boxes and counting the data item values of the records in each box.





When you partition the box to create many boxes within the box you are introducing a classification — eg a classification into time periods which we refer to as a *time series*. This could be further partitioned to introduce a classification of the statistical units (establishments) in the table according to their industry.

We could also partition the data item characteristic (number of employees) into its recorded components by dividing up the vertical axis of our data box.



Of course you could go further in classificatory detail. For example, you could *cross classify* the establishments in each industry to the geographical areas in which they are located.

You do not have to go far before you have a very large and complicated table!

Data item and statistical unit standards

The point of the *data box* illustration is that it is desirable to place limits on the number of different categories of data items and statistical units that you try to represent in statistical publications or even maintain in your statistical files.

statistical building bricks

Just as, in the building trade, there are advantages in having standard *statistical building bricks* there is a need, both for the communication of ideas and the management of statistical data, to standardise on a limited range of consistently specified definitions and classifications — both as to data items and as to the statistical units to which the data items will relate.

Even so, as later chapters will indicate, there is likely to be a somewhat daunting number of important data item and statistical unit concepts and standard classifications in respect of them. Help may be provided by special publications or electronic information facilities devoted to setting out such *meta data (*data about the data) in the form of *data dictionaries* and the like.

You will need also to become familiar with some of the other conventional jargon used in this connection. The United Nations in its various documents uses several other terms to distinguish the data item from the statistical unit to which the data item relates.

Thus:

DATA ITEM ..... STATISTICAL UNIT *or economic activity..... by an .....economic agent* *or transactions..... between .....transactors*

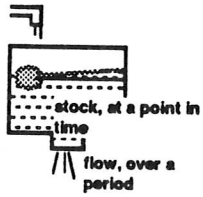
Transactions occur between *transactors* when things change ownership as a result of a sale, or change place as a result of a physical transfer.

In the first case there is an exchange of ownership of goods for a debt, with perhaps a later payment to repay the debt. In the second case, physical possession or control passes from one party to another.

There can also be transactions where only one party is involved, as when something changes physical form. The destruction of waste material would be an example, births and deaths of people or businesses are another.

transactions and statements of results are 'flows'

From these transactions, statements of results can be prepared — eg the income earned during a period is a net result of flows of receipts and expenditure. Transactions and statements of results are flows.



Data items can relate also to stocks of things or populations of people, the sum total at a point in time. The difference in the level of stocks or population at two points of time is a net flow, eg a change resulting from the difference between the inward flow and the outward flow.

It is useful to examine the way in which a statistical standards manual presents and explains its definitions of units and data item concepts and classifications.

For example, the UN System of National Accounts (SNA)[1], does this very systematically. Notice how, in its chapter on the transactions of the system (Chapter VI), it discusses each item in terms of the composition of the item, which gives its definition and discusses elements which are included and elements which are excluded.

1) United Nations Statistical Office, Studies in Methods, Series F, No 2, New York 1968.

describing units and data items

You can define a thing by describing it directly or by listing all the elements within the boundary defined. Or you can mention only those which might give trouble. Thus, it is common practice in a statistical return to describe the item in as precise language as possible and then to mention some carefully chosen examples of things which it includes and things which should be excluded.

valuation and timing of transactions

The SNA also comments on the *method of valuation* to be employed and the *point of time* at which the transaction is to be recorded. Thus a sale could have a very different value if it is valued net of discounts allowed or if it includes subsidies received. And the inclusion of a transaction in a particular period of account can depend on whether the transaction is regarded as taking place at the time the goods are *delivered,* at the time legal *ownership changes*, or at the time *payment is made.*

The SNA also introduces recommended *classifications* of an item. This helps to clarify a definition, since it is providing a definition by listing classes or subsets of the concept. So to understand a concept like *imports of goods and services* it is a good idea to examine the classification (page 117) along with the definition given in the text (page 115) or in the useful *glossary* of terms in the back of the SNA (see page 234).

Other manuals such as in the social and demographic area, will raise additional standards issues relating both to the basic elements of statistical units, data items, values and time intervals and to the choices to be made in presenting the information in standard tables. **Constantly one is reminded that taxonomic (classification) skills are critical in the development of statistical systems.**

Principles of classification

Classification involves the definition of the categories of a scheme of classification (eg the descriptions of the various dimensions in our *data box*) and the set of rules to be followed in classifying statistical units and data items into these categories.

desirable characteristics of classification systems

In a properly designed classification, the categories should be *mutually exclusive* and *exhaustive* (eg so that every statistical unit under consideration is included in one and only one class).

In addition, a classification system of any complexity will need to be *hierarchic* in structure. This means that each type of classification, whether it be an area, commodity or industry classification, is broken down from its broadest category into successively finer categories.

From the users point of view this provides a range of choice in selecting the level of summarisation appropriate to the generalisations they want to make. They can study the subject-matter in a broad way and narrow their focus where desired.

In applying such a classification, each unit is classified successively to each level of the classification starting from the broadest level. For example, a unit being classified according to the *Australian Standard Industry Classification* (ASIC) would be first classified to an industry Division, then successively to Subdivision, Group and then Class. Of course, it may not always be practicable to use the classification structure down to the finest level because of a lack of identifying information for many units.

In some instances, the statistical unit will have only one characteristic which, when correctly identified, allows it to be uniquely classified. This situation would arise in the application of classifications such as size and area. These classification systems are possibly the most simple type of classification to apply, having a relatively simple set of rules.

classifying commodities

A more complicated type of classification to apply is a commodity classification of the type used in the classification of imports and exports.

Each commodity could be classified according to a variety of criteria such as the nature of the raw material used, process of manufacture, stage of manufacture etc.

The criterion used will depend upon the use to which the classification is to be put. Each category in such a classification system needs an explicit definition so that each commodity may be uniquely classified.

kind of activity classification can be complicated

A still more complicated classification system to apply is found in systems of industry classification, such as the ASIC as they specify sets of rules which must be followed to uniquely classify the reporting unit to its appropriate category in the classification system.

Thus the problem encountered in the classification of statistical units by kind of activity is that the units often have characteristics appropriate to more than one classification category.

primary and secondary activities

This situation can be illustrated by the case of an establishment unit, such as a factory, which is engaged in more than one activity by which separate industry classes are defined. In such cases the classification rules require that the establishment be classified wholly to an industry class according to whichever activity is the main activity of the establishment.

Each ASIC class, for example, is described in terms such as the following: 'Establishments mainly engaged in manufacturing ...' (or retailing or wholesaling etc)... and then specifies a unique set of primary activities. The classification rules will go on to specify the criteria (eg value of sales, employment or value added) to be used to determine the main activity of the unit concerned.

Although we have said that the problem is a common one, it should be emphasised that in designing a classification system, particular attention is paid to the grouping of characteristics by which the units are to be classified. This should minimise the incidence of the problem, and as a result of the classification exercise, a homogeneous group of units will be classified together in the one category.

When the same set of rules is followed in each application of a standard classification, all data presented on the basis of that classification are directly comparable — ie the characteristics tabulated in each case will refer to the same group of statistical units.

Origins and levels of standardisation

One of the problems of standardisation is that it can be carried on by different authorities and with different degrees of generality. Thus each business or government agency may have its own accounting standards — each internally consistent but different from one another.

As a rule the process of standardisation is one in which standards of wide application tend to supersede those of narrower application so that the range of comparison between the accounts of businesses, government agencies and nations can be extended and the analytical power of statistical systems increased for the benefit of all concerned.

how to extend the scope of standards?

Common standards between agencies collecting economic and social data are not arrived at easily. Typically a collection or compilation of statistics is undertaken to meet some immediate special need for information. Without some prompting from a central statistical office the need to associate the collection with other statistical series may not be foreseen when the programme is initiated.

Once a programme is established, its definitions and classifications are very difficult to change — a lot of work has gone into it and the people who have done the work may have a vested interest in keeping it that way. And of course, unless it is possible to revise figures for previous periods, any change involves a break in the statistical time series, to which some users may well object.

Thus the introduction of national and international statistical standards is likely to require a sacrifice of comparability over time in order to obtain a gain in comparability between different subject matter, geographical regions and countries.

When agencies are faced with the task of devising definitions and classifications or are reluctant to change to new standards of wider application, **the existence of an agreed international standard can be decisive**. In addition, international programmes, such as the 1973, 1983 and 1990 *World Programme of Industrial Statistics*, help to overcome inertias and encourage statistical offices to move ahead in a concerted way.

The process of determining international standards

There are a number of international statistical organisations and forums such as the *International Statistical Institute*, the *Inter-American Statistical Institute* and the *Conference of British Commonwealth Statisticians*. But the most important since the 1939-1945 World War has been the United Nations organisation with its *Statistical Commission*, its Regional Economic Commissions, such as the *Economic and Social Council for Asia and the Pacific* (ESCAP) and its Specialised Agencies.

The initiative in reviewing statistical standards comes generally from the *Statistical Commission of the United Nations*. The Commission is composed of representatives of national statistical offices who meet every two years. Its job is to consider such problems as promoting the development of national statistics and improving their comparability, and the coordination of the statistical work of Specialised Agencies such as the *International Labour Office* (ILO) and the *International Monetary Fund* (IMF).

the initial drafting

The process of reviewing a standards manual would typically start with a request from the *Statistical Commission* to the *UN Statistical Office* (which is a permanent office within the UN Secretariat).

This office would draft a proposal with the help of an expert group appointed by the UN, and advice from relevant UN specialised agencies and international professional organisations. (Increasingly, the carriage of the review may be entrusted to such bodies as *OECD, Eurostat*, the *International Monetary Fund* and the *International Labour Office*).

The draft proposal would be referred to the various Central Statistical Offices of national governments for comments. On the basis of these comments the proposal may be redrafted and put to the next meeting of the Statistical Commission. Before doing that in the case of important proposals (such as the review of a manual) the proposal may be discussed and commented on by regional working groups of experts from statistical offices in each region and may come before a regional conference of statisticians, which would report on the matter to its regional economic commission.

After the *Statistical Commission* accepts the final proposal, it still has to go before the *Economic and Social Council* and finally its recommendation goes to the national governments for them to implement.

Of course a major manual such as the United Nations *System of National Accounts* (SNA) or *Towards a System of Social and Demographic Statistics* (SSDS) may go through several rounds of revision over a period of several years and the implementation and further elaboration of the system may go on for years after that.

What should at least be clear is that these standards are not something thrust on government statisticians. They are the result of a great deal of thought and effort in which the statisticians themselves participate in order to fix standards which will be generally accepted as guidelines for all countries to follow.

Establishing standards in practice

The publication of agreed standards on conceptual and operational aspects of national statistical systems in different fields is not the end of the process. It will take some time and effort at international levels before the standards are generally adopted.

For example, pressure on Asian and Pacific statistical agencies is exerted by the statistical office of ESCAP in the course of its publication programme which brings together data from the countries of the region which are requested to report as far as possible on the basis of the international standard. Agencies like the South Pacific Commission and the International Monetary Fund exert similar pressure via their reporting requirements.

The central statistical offices within each country in turn try to publish consistent national statistics and exert a pressure on units in their system.

Of course, as new standards are put into practice they may be found to be in need of modification in different countries as circumstances vary. Over a period it may become apparent that the international or national standards need to be reviewed and revised. **For example, revision 4 of the SNA is likely to follow some 25 years after the publication of revision 3 in 1968.**

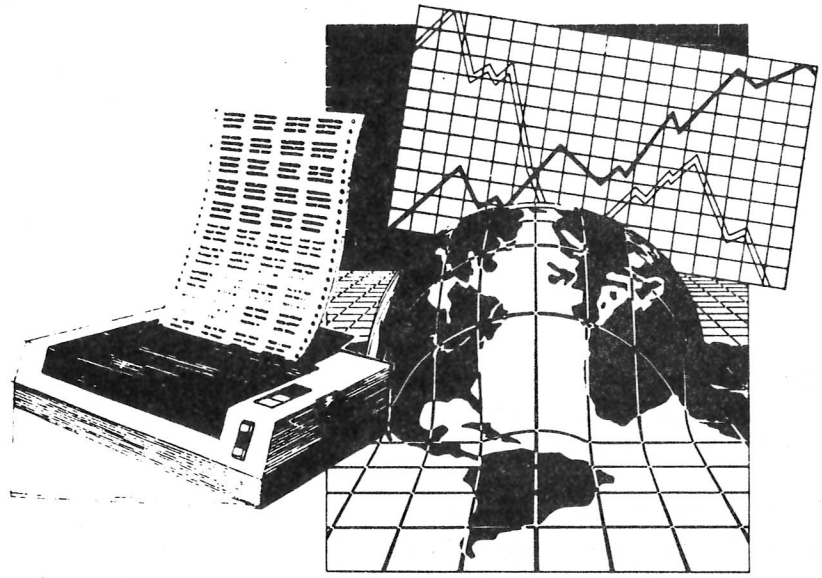
In principle and in practice

It is important to appreciate that standards in statistics, like standards in any other field, are objectives which are not always fully realizable in practice. Thus one often finds in discussion of standards that a manual may say *"in principle one should do this, in practice it may be necessary to do that"...* Similarly, one may speak of the *theoretical concept* on the one hand and the *operational concept* on the other. A good statistician will always try to specify what system of standards has been followed and will indicate any significant departures from this.

Some countries which accept a set of standards like the SNA in principle may never be able to implement its standards fully. This should not be seen as a fault of the standards themselves.

> *Thus it was argued in the case of Papua New Guinea that 'accepting the SNA as the framework for developing economic statistics means no more than accepting the basic structure of the system and using its standards as to concepts, terminology, forms of presentation etc as (at least) the point of departure. It does not imply commitment to any more sophisticated statistical presentations than are needed or are feasible in the light of the country's circumstances. It can serve as a valuable framework within which to consider future development and standards to follow'.*
>
> *Eighth Conference of Commonwealth Statisticians, Bridgetown 1975 contributed papers p. 117*

Thus the adoption of common conceptual standards is essentially a process of improving *communication.* It is not that there can be only one correct definition of a thing. **It is just a lot more convenient if everyone accepts the same definition and it becomes part of their working vocabulary.**



a point of departure

The international recommendations may not meet all the local and specialised needs for statistics but it is always possible to adapt them so as to be able to produce statistics on an internationally comparable basis while also meeting the special needs.

In doing this it is very desirable to use the international recommendations as the point of reference from which one should depart only when there is good reason to do so.

Comprehensive standards frameworks

The SNA illustrates comprehensive standards systems to which less general international standards should be seen as subordinate.

Thus the SNA is the most comprehensive set of recommendations relating to economic statistics and as such can be regarded as an ultimate authority in the development of standards for statistics for narrower economic fields, like balance of payments statistics.

The rationale of the definitions and classifications of data items and statistical units is made plainer in the SNA because the relationships between them are shown in the accounts and tables for each sector.

Its general accounting scheme forces consistency between items, both as a theoretical system of ideas and as a system for presenting collected and estimated data.

A whole range of UN Statistical Commission manuals maintain consistency with the SNA, while spelling things out in greater detail and providing for rearrangements of tables to meet specialised needs. This carries over into social and demographic fields where, for example, the SNA-based industry classification (*International Standard Industrial Classification of all Economic Activities* — ISIC) will enter into such frameworks as *Towards a System of Social and Demographic Statistics* (SSDS).

To sum up

The development of a national statistical system proceeds as much through literary processes of defining and classifying as through numerical processes of collecting and manipulating figures.

Standard definitions and classifications of statistical units and data items are essential elements of a national and international statistical system. Beyond this, a whole range of standard practices, relating to ways in which statistics are collected, processed and published, need to be developed as part of the working philosophy of a government statistical service. Such standards are developed at an international level with a strong lead from the UN Statistical Commission.

The keynote of all this is communication. Common understanding between suppliers, producers and users of statistics of what is required to be measured and how it has been measured is as important as the statistical values obtained.

PART 2: PROCESSES

CHAPTER 5
A national survey operation — what's involved?

Steps in the process of producing and using statistics
Technological changes in data processing
Survey design and work scheduling elements
Data collection system elements
Data processing system elements
Processing control
Dissemination system — final review and publication
A word for the users

CHAPTER 6
Continuing sequences of enquiries and a rolling development cycle

The one-off operation
How much? how often?
'Wheels within wheels' reduce the time lag problem
The census/survey choices
Relationships between periodical census and annual or more frequent enquiries
Preliminary results and rolling revisions
To sum up

CHAPTER 7
Integrating enquiries across subject matter fields

How does it happen?
What are the elements of integration?
What would we be looking for in a national economic census and survey system?
How integrated can you get?
How widely can we integrate?
Theoretical frameworks
Standardisation of concepts
Vertical and horizontal integration
Integrated collections
Instruments for integration
Wider still?

CHAPTER 8
Managing the stock of data — the data base approach

Elements of the data base approach
What goes into a statistical data base?
From the users' perspective
How is the statistical production process affected?
Confidentiality
Access rights and security
To sum up

CHAPTER 9

Using the data stock in analysis, synthesis, problem solving and system feedback

CHAPTER 5

A NATIONAL SURVEY OPERATION — WHAT'S INVOLVED?

What goes on in a national statistical office? Part 2 of this 'guided tour' concentrates on the processes involved in producing and using official statistics. It will try to bring out the special features of large-scale national statistical systems, starting with the most straight forward case of a one-time survey before going on to more complex continuing and interrelated programmes.

In this chapter we consider how a national survey is conducted and how a large data set is assembled from this and processed to meet pre-defined information requirements, proceeding to timetable to produce a set of statistical tables. In later chapters we will introduce the data base idea in which the collected data from many such sources may be stored in the one integrated file system in a state of readiness for retrieval and processing as and when needed by a wide range of users.

At this point it is worth noting that such comprehensive data bases are still dependent on statistical programmes and have to be planned with purposes clearly defined. The purpose may be more general than in the case of a specific survey, but it still has to be specified and evaluated in developing the most efficient system on cost/benefit lines.

Steps in the process of producing and using statistics

Logical steps in producing and using statistics to help solve a problem are likely to include the following:

(i) **recognise a problem** requiring information for solution (eg existence of poverty in a certain area);

(ii) **define the conceptual boundaries** of the problem;

(iii) **design a statistical model** (identifying units to be measured and their relationships — e.g. a model indicating the elements of income determination);

(iv) **define the statistical output** required. Then determine the data that can be obtained to provide this output (or get as close to it as possible), as well as the manipulation of the input data required to produce the output;

(v) **design and test** a system for collecting the data required;

(vi) **collect and check data** (which involves the prior listing or mapping of statistical units to be surveyed)

(vii) **manipulate the data** (eg systematically arrange the data and compute aggregates, percentages etc);

(viii) **present the data** (eg in tables and graphs);

(ix) **analyse the data** (this may lead to further manipulation and fresh presentation, or a return to the first step with a new view of the problem);

(x) **interpret the data** (reach conclusions in relation to the problem);

(xi) **determine actions to solve the problem** (this may lead to a fresh cycle, eg to solve the problem of ensuring that results of the action to be taken are continuously monitored)

Most of the time and resources of government agencies producing statistics is absorbed in steps (ii) to (ix) of the process illustrated. But the other steps have to be taken to complete the cycle and the government statistical agencies must be involved in them, at least to the extent necessary to ensure that the statistics produced satisfy the users' need.

the plot thickens!
— systems of surveys

This sequence of steps readily describes what happens when a one-time survey is undertaken to meet the requirements of particular users who have a clear specific purpose in mind. It can be more difficult to recognise the process when examining the large-scale general-purpose interrelated collections of the kind undertaken by central statistical services. Although the same process is involved in the continuous development of the collection or collections, it can perhaps be better viewed as a process of developing an information system to service a wide range of present and potential user needs.

Some of these operations are manual, but most today involve computers in one way or another. As a result, the character of the statistical production process has changed considerably. This has some important implications for the user of statistics.

Technological changes in data processing

In most countries, certainly in developed countries in the 1960's and 1970's, there has been a quiet revolution in technique which has forced some return to fundamentals, as many long standing operations have become redundant or have been transferred from one kind of data processing specialist to another.

The two major (and related) influences in this are:

- the development of central lists of statistical units (*registers or dlrectorles*) which are maintained on a continuing basis to provide a common population listing (eg of employers or households) for a whole range of statistical censuses and surveys;

The black box!

- the advent of *computers*, and the transfer of large blocks of formerly manual operations into the *black box* realm where operations take place automatically once appropriate programs are designed and tested.

The regular operations of collecting, editing, manipulating, tabulating and analysing data, from the point of view of the subject matter statisticians, have reduced largely to preparing a 'clean' input of returns and to handling the output through the final scrutiny stage to the point of clearing manuscript for publication. **The computer systems which make this possible are likely to incorporate frequent interaction with the central register system as an active element in all phases of both data collection and data processing.**

implications for survey planning and operations?

These influences have meant that much of the tedium of the work of producing statistics has been removed, but the intellectual element of the task has simply shifted in timing. Much more of it is now done in advance — for example, in designing, constructing and maintaining central registers to serve as the basis for a number of operations in a number of statistical programmes.

Similarly many of the processing decisions made in the course of a census or survey operation are now programmed and tested well in advance so that the decisions are now predetermined for a wide range of anticipated situations.

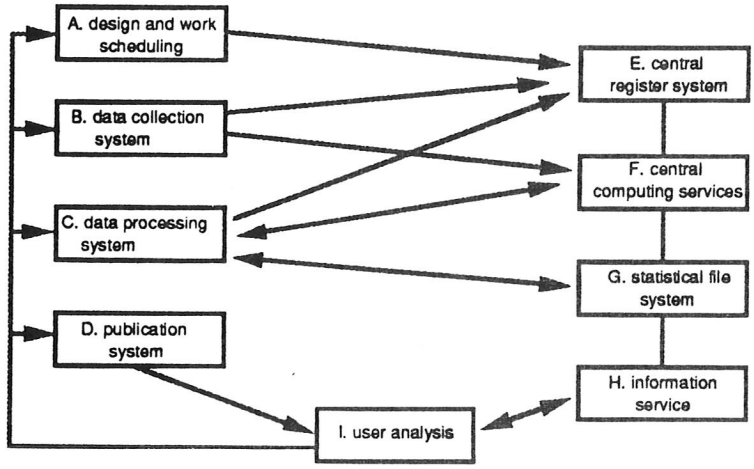
Because so many of the formerly unwritten rules and practices have had to be made explicit and be subjected to intensive critical examination, the result has been that the intellectual content of the task has increased in the planning phase, both with the need to translate and transform the process in this way and with the possibility of doing a better quality job. At the same time this task has been simplified somewhat by the availability of *generalised programs* which have reduced the need to write programs on a *one off* basis for every census and survey.

new output facilities

The introduction of computers and electronic data storage, providing fast access to the data plus the facility to generate special statistical compilations economically at short order, has added an additional dimension to the publication end of the statistical production process.

In effect, publication is no longer a real end point. There is rather an information service, producing user-defined tabulations as required from a data base of computer files continuously updated and readily accessible, often in computer readable form.

The boundaries of this present discussion can thus be defined in terms of the following representation of the statistical production process for a large scale repetitive enquiry:



The elements A to D summarise the operations of the kind of independent statistical programme common in the past and described in the text books, which tends to develop into the large scale interconnected statistical programmes in which E, F, G, and H provide major support elements today.

Survey design and work scheduling elements

In this phase of the process, basic decisions must be taken in respect of all the stages to follow, including the specification of the output of tabulations required.

Thus it is often not until the users specify the tables they will require, that their expectations of the information they will get from the survey become clear. The specification of the output requirements in the form of draft tabulations helps to ensure that the questionnaire design, the intermediate processes of editing and coding and the production of tabulations are consistent with the survey objectives.

who to survey?

The design work will include the firming up of the SCOPE or conceptual boundaries of the enquiry (ie the defined range of statistical units in respect of which data are to be collected). Thus it may be decided to survey all the establishments in a particular range of the industry classification, or all households having certain characteristics.

It will also require decisions about the COVERAGE of the survey, ie whether an effort will be made to update maps for household surveys or the central register's list of employers, to ensure that all units in the defined scope will be covered in the population actually surveyed.

about what?

The scope and coverage specifications thus define the POPULATION to be surveyed. Depending on decisions about the DATA CONTENT and levels of ACCURACY required, and the RESOURCES available to conduct the enquiry, an appropriate survey design may then be developed.

The word 'survey' is used here in the general sense as including a full census, ie a 100% sample. By the same token, a sample survey is likely to cover the large size units in a population (ie it will have a completeiy enumerated 'census' component).

how to survey?

If it is decided to use a sample survey, a variety of options about the most efficient method of SAMPLING may have to be considered along with the consideration of the data collection method (which may involve obtaining lists of statistical units or maps of localities from which to select the sample).

how thoroughly?

In general, a sample survey will provide results at a fraction of the cost and more quickly than will a complete enumeration, but at the expense usually of geographical or other classificatory detail. The introduction of an additional source of error may be compensated for, perhaps, by better quality control of input data.

information quality/cost options

The problem for the survey designer is then to determine the size and structure of a sample which will provide estimates with the required degree of precision economically. Alternatively the survey designer might design a sample which will give the most reliable results for the money available.

Of course there can be a wide range of options as to information quality and cost of collection.

Suppose we wish to establish the fuel consumption of the local bus fleet.

Firstly, one could ask a fellow worker (who in most cases probably would have little idea).

Secondly, one might ask an expert, that is, someone within the transport section who should have a nodding acquaintance with these sorts of facts.

Thirdly, one might ask a number of people or see a few of the items; that is, to do some of the measurement necessary to obtain these statistics. This would involve using some of the buses available to do some quantitative measurements on their fuel consumption.

Fourthly, one could make these measurements on all the buses available at the present time.

Fifthly, one might carry out a properly designed sample operation to make, on a sample of all buses currently in use, measurements of fuel consumption to enable unbiased estimates to be made of the required statistics.

Sixthly, at the other end of the scale, one might measure continuously the fuel consumption of all buses at all times.

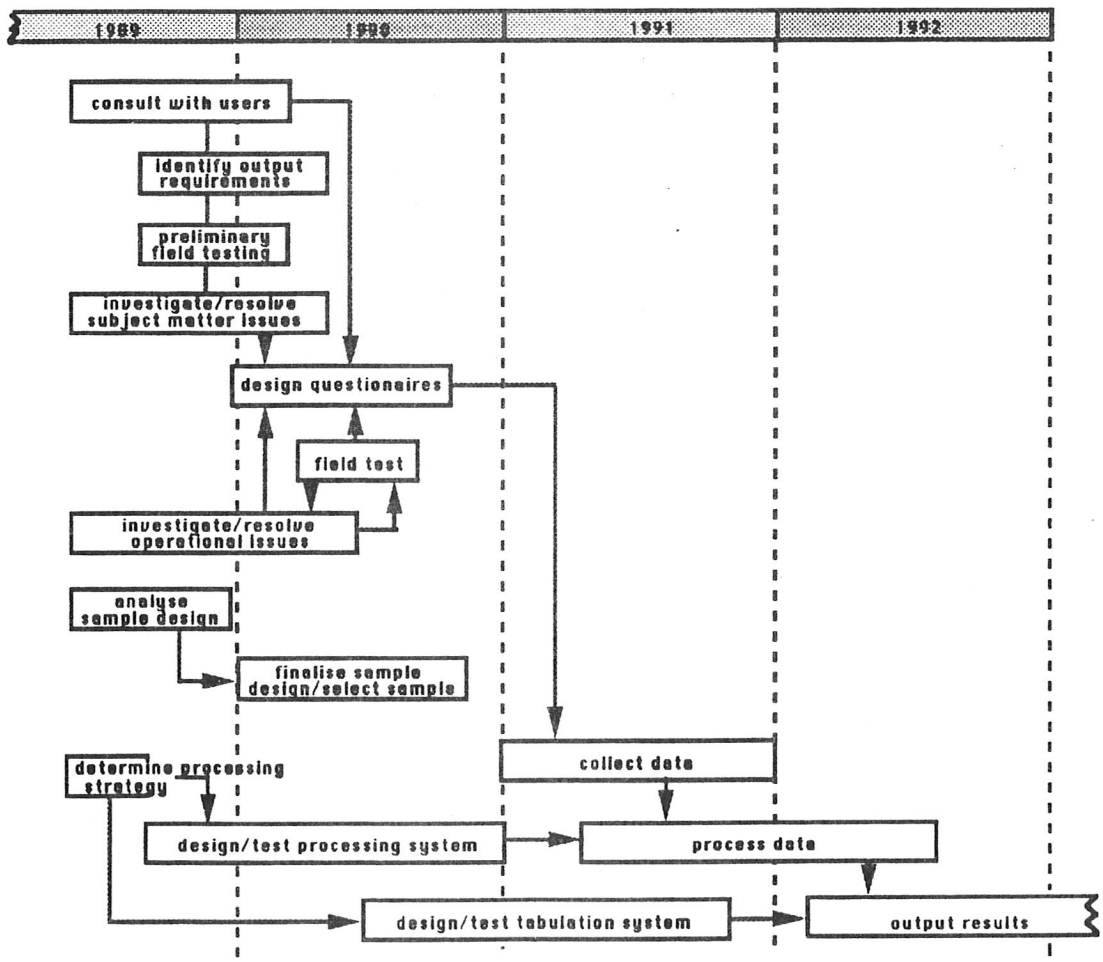
Throughout this spectrum of methods, the cost of collection increases; experts cost money, properly designed samples are more expensive, and the complete enumeration described would be very costly indeed.

The DESIGN and TESTING of the QUESTIONNAIRE is crucial to the survey operation and again there are trade offs to be considered as to the detail to be collected and the cost and precision of the operation. On the one hand, there is the list of items which the statistical model prescribes as necessary or desirable to collect. On the other hand there are more-or-less equivalent items which are judged to be readily available from respondents.

how to ask the questions?

The data may be available, but the problem may be in defining items in words that are succinct and clear. Care has to be taken to ensure that the number and complexity of the questions asked will not produce a resistance on the part of the survey respondents resulting in inaccurate reporting or a sufficiently high non-response rate to compromise the representativeness of the survey.

SAMPLE WORK SCHEDULE FOR A ONE-OFF HOUSEHOLD EXPENDITURE SURVEY



At the same time the designer must be sure that, with whatever manipulation is necessary, the data collected will provide the information output that users need.

when?

The SCHEDULING of the operation involves establishing a timetable for all elements of the task and commiting resources to meet the timetable.

Particular attention has to be given to critical points in the timing to ensure (eg, in the case of interviewer-based collections) that enumerators are recruited and trained in time and that the collection forms and other material are in their hands, fully tested and proven in earlier pilot surveys.

Data collection system elements

data gathering options

In this phase the procedures and the resources of personnel and material have to be ready in time for the data collection operation. The possibilities include various combinations of physical OBSERVATION (eg 'block listing' of houses) or measurement (this would be rare in social surveys on a national scale); personal INTERVIEW (requiring trained field or telephone interviewers — perhaps facilitated by computer automated telephone interview systems); MAIL enquiry (with possible field or telephone follow up); REGISTRATION processes (eg the requirement to register births and deaths) and extraction from ADMINISTRATIVE RECORDS (eg from the import and export warrants which are required by Customs' authorities).

Commonly data are collected from households by field interview while surveys of businesses and other organisations are generally by mail (although field interviewers may be more effective in collecting from small businesses, especially in developing countries where address listings and mail deliveries are unreliable).

Extraction from registration and other government administrative records is a valuable source of information both for households and for business and other organisations. It has the advantage that collection, data entry and editing is part of the administrative process, so that the direct cost to the statistical agency may be low. However, this can raise problems for the statistician when (as is often the case) he has little or no control over the design and the quality of initial processing of the records and may not have direct access to them.

Data collection normally takes longer in the case of a *mail collection* than in the case of a *field survey*. In the case of *household interview* surveys, if the intention is to collect information relating to a particular point or period of time, the effort is usually made to concentrate enquiries into a short data collection period so as to reduce the risk of errors of recall.

The problem of recall generally will not arise in mail collections in respect of the records of businesses or other organisations. Nevertheless, if the statistics are to be timely, it is necessary to have as short a collection period as practicable.

In mail collections, collection control procedures must ensure that, after the expiry of the date for the form's return, a tight reminder programme is followed through until either the form is received or a decision is made to either estimate for the missing return or to omit it from the survey.

Data processing system elements

Returning to the diagram on page 42, the basic operations in the *data processing system* contained in **C** are essentially the same with or without the development of the facilities indicated by **E**, **F**, **G** and **H**.

It could be said, of course, that the total system A-I is a 'data processing system' — not just C. But perhaps the term 'statistical information system' is more expressive of the overall task of 'data capture' (data collection), 'data processing' and 'data delivery' (publication and information retrieval from statistical files).

Firstly, the collected data have to be checked at the level of the individual report or record, **secondly** the data in the records have to be converted into the statistical compilations required, **thirdly** these have to be checked and then, **finally**, they are made available.

Any attempt to elaborate this description of the data processing function very quickly introduces the complication that each of these four elements can be subdivided into many tasks which may be very different in character and in time sequence, depending on the size, quality and complexity of the record input and the required statistical output.

Even when these are assumed to be broadly the same, significant differences can occur in the flows of information within the statistical office according to the various degrees of manual and computer processing introduced.

Before pursuing this, it will be desirable to discuss some of the operations which will occur in one form or another in any system of survey data processing in order to at least recognise the terminology which is used to describe elements of these systems.

converting to processing medium

DATA 'PREPARATION', DATA 'ENTRY' OR DATA 'CAPTURE'. These terms describe the task of converting the data from the collection record (eg the statistical return) into the data processing medium eg. disk or magnetic tape.

Data transcription to a processing medium is also usually necessary in manual processes — in this case it involves entering details extracted from the returns onto standardised working sheets for convenience in the subsequent checking and summarisation processes. But it is possible with simple forms to sort the forms into the required categories and add them directly, entering totals on blank forms.

With the almost universal use of computers, manual data preparation is likely to be limited to screening the returns to decide whether they are *clean* enough to be converted into the form which the computer can read or whether they should be queried with the respondent. Generally, the object is to get the data into computer-readable form as early as possible in order to do much of this checking by computer.

checking the data input

EDITING AND CODING. As viewed here, editing involves checking the quality of the information reported on questionnaires and entered into the computer a) to detect missing, inconsistent or incorrectly reported or entered data, and b) to take corrective action where required.

Particularly with the introduction of computers as part of the processing system, it is useful to distinguish two stages of editing — the first being mentioned above as associated with data preparation and the second being undertaken with the help of the computer.

First stage editing will involve

(i) PHYSICAL CHECKS to ensure that, for example enough questions have been answered to warrant further processing and

(ii) putting the recorded data in the form required for entry, such as by rounding a number in a standardised way (eg to the nearest $1000) or by striking off unwanted digits.

(iii) The further step of ARITHMETIC AND LOGICAL EDITING (eg to see that one reported figure is consistent with another on the return or on previous returns for the same unit) is likely not to be pursued very far by manual means where computer editing is a viable alternative.

assigning classification codes

Associated with the first stage editing in the data preparation phase is likely to be some coding. CODING is the transformation of information into numerical or other condensed form. For example it may consist of assigning numbers or letters to items on the return to

permit ready classification of the statistical unit by such characteristics as industry, location, size and type of ownership.

Assigning size codes, based on numbers of persons employed, transforms one set of figures to another and simpler set for tabulation purposes. *Geographic coding,* on the other hand, transforms the names of regions, local government areas and cities to a compact numerical form.

automatic coding

This is a process that changes character with application of computers and the availability of current registers of statistical units. If other means than the survey questionnaire have been used to keep the register file up to date, such classifications as these should already be determined on the file and codes do not have to be assigned manually during the course of the survey. However the computer may be used to check and correct the classification.

For example, it may compare the sales data supplied for the different commodities produced by the establishment and find that the establishment's predominant kind of activity has changed, or it might, as a matter of course, classify the establishment according to the size of its employment.

Any coding done for computer output can be limited to the most detailed level of a classification — thus, if the code of the street section in which a statistical unit is located is punched in the unit record, the computer can translate this into the local government area, region or State as necessary for different tabulations by consulting a *look up* table kept on a computer file.

Thus clerical coding is generally limited to the classification of records based on information which requires reference to a variety of textual information (eg name and address and description of a business's activities). The coding of products, occupations etc may also be necessary if precoded lines were not provided on the collection schedule and product titles had to be written in.

As it becomes possible to enter more information of this kind into the computer, even the more complex classifications can be left to the computer system. Where the assignment of a code depends upon a computation and/or comparison of relative magnitudes, this can be done more accurately and economically by the computer. This is likely to be the case with industry coding, for example.

automatic error identification and correction

Editing of data, with or without a computer, can take many and varied forms depending on the nature of the enquiry. Systems for identifying and correcting errors in the returns, as received or entered, tend to be improved with experience in on-going statistical programs.

It is also possible (particularly with a computer) to compare the values in any one record with the general run of values for a batch of data in order to isolate exceptional cases for query.

Of course this is where human judgement and experience enters into the process — exceptional cases can be quite genuine and other information on the record (or perhaps a clerical record of a previous investigation) will indicate that the unusual report is reasonable. There has to be a balance, in specifying tolerances for the isolation of queries, that ensures that the number of cases identified for query is not too large to handle with the staff available while at the same time ensuring overall data quality. These judgments should also consider the proportion of queries which result in changes to the data.

It is often possible to introduce procedures to correct data without time consuming reference back to the informants. For example, this is not necessary when missing data, that is, data which should be there, can be estimated with confidence from other sections of the record or batch and inserted automatically.

This operation can be done for individual units or for low levels of aggregated data. For the published aggregates the results will be the same by either method. However, correction of the individual record is better if it is likely that the records may be used again for microanalysis.

QUERY ACTION AND AMENDMENTS. Query action involves withdrawing the record from the processing flow while reference is made back to the respondent, or some other information, and then (frequently) an amendment of the return or other record and its reintroduction into the processing flow.

output preparation

TABULATION AND REPORT GENERATION. One of the main statistical functions for many collections after editing, amending and transferring data to worksheets or other processing media is to accumulate, tabulate and report the information in tabular form. This is a long and tedious process in a manual system, involving transcriptions to summary work sheets, providing for the different classifications and cross classifications required.

With the use of the computer, the table layouts are specified, programmed and quickly produced. The subject matter statistician's task is then concentrated on scrutinising the tabulations and applying various quality checks in the process of preparing a publication.

In practice there are likely to be program and computer *hang ups* to contend with, but when the *bugs* are removed in testing prior to production runs, the tabulations can be produced quickly and painlessly. With the introduction of generalised systems to produce the accumulations in tabulation cells (*table generators*) and to

assemble these into readable, standard tabular formats (*report generators*) even the specification of the job is now a relatively simple task.

In this phase, other computations besides accumulation may be carried out, such as the expansion of sample aggregates in accordance with the sampling plan, calculation of variances, etc. The tabulations may be tailored closely to correspond to the format of tables for publication and can include various analytical ratios such as percentage changes in a time series of aggregates.

Processing control

What has been described are the essential steps common to the data processing of a large census or survey such as a census of industrial establishments. The form and sequence in which these tasks are organised is likely to vary considerably from survey to survey.

Depending on the complexity of the process, there will also be a more or less complex coordinating function of maintaining administrative control over the operation and of supplying management with continuous information on the progress of the job.

Thus it is necessary to control the movement of all documents into, out of, and within the various sections involved in the operation, to ensure that all records pass through the successive stages of processing without any unnecessary delays and that the data processed are complete and accurate. **The computer systems designed to process a survey will generally provide all this information as a by-product.**

Dissemination system — final review and publication

This function is treated separately from that of tabulation and report generation because (particularly with computer processing) it involves a different group of specialists with special editorial skills and experience with publication processes. It also involves the subject matter specialists in close interaction.

The important timing consideration is usually the publication of preliminary output in as short a time as possible, so that the publication phase can be thought of as consisting of a PRELIMINARY RESULTS SYSTEM and a MAIN TABULATIONS SYSTEM.

Deadlines are set to complete the preliminary results but production of main tabulations can be expected to continue for some time and to be followed by many special purpose tabulations, requests for which arise from time to time.

Planning of at least the broad content of the statistical tables should always be undertaken in the preliminary design phase of the survey. There will still be decisions to be made subsequently on the exact final form of the tables and the accompanying material such as text, including technical descriptions and footnotes. These decisions will be taken within the framework of publication standards which become established in the larger agencies.

output scrutiny

Decisions as to whether the data are sound enough for publication cannot be made until the data are finally examined along with the evaluation of the editing programme and information from other sources (eg national accounts estimates) which might provide an independent check on the final aggregates.

The final scrutiny of publication manuscript after careful check reading and check adding of all figures and text will include a logical edit for internal consistency — eg to check that totals for a given item appearing in several tables are the same — and for consistency with any relevant external data.

To some extent this may be duplicating earlier checks of tabulated totals, but it is necessary nonetheless. The review of data can be helped by the computer calculating ratios and other review comparisons — eg ratios such as male/female births, or the average annual wages and salaries and cost of materials as a percentage of the value of an industry's output. An item by item comparison of the percentage changes from the previous survey may be illuminating.

disclosure checks

An important part of the review of tabulations prepared for publication is that of examining publication cells to ensure that there are sufficient statistical units represented in the cell to be able to publish without risk of disclosing the figures for any individual respondent. A disclosure analysis can be programmed on the computer but, at some level, judgement is likely to be required — eg in combining cells to avoid disclosure. In a table with horizontal and/or vertical totals, at least one more cell than the *disclosure cells* must be suppressed to prevent the offending cell being derived by subtraction.

This problem of *residual* or *consequential* disclosure extends to the possible comparison of one table with another table in the same publication or with tables published at different times. It can arise in any overlapping sets, as when data in special geographical or industrial classifications made available on request might be compared with already published data for slightly different groupings.

Once the preliminary and final results have been cleared for publication they may be disseminated through a variety of media. In addition to the printed publications tailored to the requirements of different user *markets*, statistical information is accessible through computer print-out, microfiche, magnetic tapes and discs and electronic transfer on-line from public-use data bases, sometimes via facilities maintained outside the government statistical agency.

Such facilities open up data to the public for analysis as and when required, within the constraints of the levels of aggregation in the records accessible to the public.

In fact, associated with the technological advances, there has been something of a revolutionary change in statistical dissemination philosophy in the 90's as countries have moved from an 'input-processing, standard product' orientation to a much more flexible output 'market-oriented' stance. On the one hand, new systems are enabling statistical services to respond dynamically to client needs. On the other hand, services are being charged for and much of the development is being driven by market demand.

A word for the users

What has been described is the kind of statistical production process involved in conducting major censuses and surveys — direct collections which form the data base for analytical extensions of the system (including the regular official compilations of national accounts, price indexes and the like). Users who seek to influence the development of the system must be aware of the design lead times involved with computer-based systems and especially as separate enquiries are combined into much larger integrated systems. The crucial stage in which users can influence the design of a survey is in the design and programming phase.

CHAPTER 6

CONTINUING SEQUENCES OF ENQUIRIES AND A ROLLING DEVELOPMENT CYCLE

The initial discussion of the characteristic processes of a national statistical system introduced the various functional elements as they might be involved in a one-off national survey. This chapter elaborates on this model in order to indicate how continuing collections are managed and the issues that arise in their development.

The one-off operation

One-off implies a collection undertaken to supply information that was not previously available and which is now required to meet the requirements of a particular client or group of clients.

The process starts with a careful assessment of the information requirements and their translation into the specifications of the collection. The designers should aim to anticipate the needs of potential users of the data as well as the needs of the prime customer initiating the project.

Basic questions will relate to cost benefit considerations — how much detail should be collected? by mail? by interview? Will a sample rather than a census enable the requirements to be met with sufficient reliability and detail?

In all these issues the designer of the one-off survey has a *clean slate* in putting together a design which will be an optimum solution, given the constraints of time, resources and so on.

Is it really a 'one-off'?

Of course, there will be surveys which are widely spaced and irregular enough to be regarded as one-off, but which nevertheless build on the experience of previous surveys and should replicate them to the extent that the information requirements include comparability over time.

has it been done before?

Particularly when there is a possibility of the project being repeated at some time in the future, it is standard practice to conclude with an evaluation of all aspects of the project, recording information that will help to ensure that mistakes are not repeated and that experience accumulates, notwithstanding turnover in personnel.

Even with the strictly one-off surveys, the designer's *slate* is unlikely to be completely clean at the start. It is probable that the survey will not be an unique one — a similar one may have been conducted in another country or it may have some aspect in common with other projects that have been undertaken previously by the agency, relating to the same population of statistical units or to the same kinds of information to be collected.

The survey designer will need to take account of any such relevant experience and, in fact, such professional interchange is characteristic of national statistical systems.

international influences

Sometimes the professional interchange is stirred along by agencies promoting international statistical standards, such as the *UN Statistical Office*, in the interests of ensuring that as much detailed and comparable data is available as possible for a particular field for all countries. With the additional stimulus of technical assistance, most countries have responded to invitations to participate in, for example, the decennial *World Programme of Industrial Statistics* and decennial Population Census rounds.

How much? How often?

As new statistical information is made available to users, the supply of the information seldom reduces the demand. It tends rather to generate additional demand for further detail, greater precision and for regular monitoring of the activity in question.

time horizons of planners

The government statistical service's response may be to establish regular collections and compilations which provide time series geared to the time horizons of the government planners and managers which the NSS is primarily intended to serve. Whether this happens should depend on whether the cost of establishing a time series is justified by the value of the information thereby provided.

Thus *medium term planning* — eg a 5 year development plan — will call for detailed information on the structure of the economy. *Long term perspective planning*, looking at trends a decade or more into the future will need to be related to historical series running back many years — such as broad national accounting data for the main sectors of the economy. *Short term planning and management*, such as might be undertaken annually or quarterly, will necessarily be based on a much more broad brush monitoring of change, with an emphasis on economic variables which can be influenced in the short run, such as production, income, employment and capital expenditure. There will be less interest in allocation of resources, which are matters rather of medium term and long run policy.

If the emphasis is on *regional planning* the planners will be most concerned with resource allocation and usually will prefer small area detail of the structure of production at moderate frequency (say 5 yearly) to broader aggregates at frequent intervals.

Of course, as a country's planning and management becomes more sophisticated, the planners will want *both detail and frequency*. There will be planning at both

national and regional levels and continuous monthly and quarterly economic management will be nested within 5 year detailed plans and long term broad perspective plans. Statistical agencies will be wise to develop their information systems with such ultimate requirements in mind.

'Wheels within wheels' reduce the time lag problem

data collection cycles within longer estimation/analysis cycles

The planner of statistical systems needs also to allow for the fact that there are necessarily time-lags between the collection of data for a given time period and the delivery of the initial statistics for that period. One consequence of this fact is that rounds of economic planning and analysis of a particular frequency may generate a demand for rounds of data collection of greater frequency. Thus:

- *historical analysis* needs to be, for example, at least 5 yearly for 5 year development plans, annual for annual budget planning, quarterly for continuous economic management;

- *projection* of data should be based on as current an historical series as possible. This calls for a minimum of lag in estimation for the current period — eg current quarter economic indicator estimates in the following month, current annual estimates in the following quarter;

- *current estimates* may be derived analytically by the statistical office from directly collected data for part only of the current period — eg for 1 or 2 months of the current quarter in a quarterly national accounts bulletin, 2 or 3 quarters of the year currently under review in the annual national accounts publication.

timely annual analysis needs quarterly surveys

This suggests the broad rule of thumb that timely planning and analysis at annual intervals requires basic data for quarterly or half yearly periods (as well as data for annual periods). Thus the provision of a major data collection at a given frequency may lead to a demand for some less detailed but more frequent or timely *indicator* series to enable current period estimates to be made with minimal time lag.

trade offs between timeliness, detail and precision

When economic planners and managers elect to plan and review more frequently than before, they must be prepared to sacrifice detail and precision in order to make it possible for more frequent and timely statistical estimates to be provided.

To start with, they must make do with measures of key items only, and may have to accept estimates of percentage change rather than firm aggregate measures. **But of course the supply of statistics almost**

invariably stimulates further demand for something more detailed and reliable than the broad indicators which are initially provided.

What has been making the pace in many countries is the use of econometric models — these are constantly being refined and so are constantly challenging the statistician to provide more elaborate and better integrated current data for forecasting.

the statistical design gestation period can be elephantine!

The problem which arises for the statistician is that demands involving major changes or elaboration in data reporting and processing systems cannot normally be met at short notice.

Statisticians must therefore try to anticipate the demands which are likely to be made on them. For example, if they are planning a central register of employers to be used for proposed annual censuses and surveys, they would be wise to also provide in its design for the possibility that it will be used later for more frequent than annual surveys.

time series continuity

At the same time the urge to improve and elaborate a survey system must not compromise the reliability of estimates of movement. Generally the main purpose of running continuing systems is to monitor change. This calls for (eg) producing overlapping estimates if the population coverage is increased, and *bridging* of some type when there are significant changes to item definitions or other reporting specifications or to the sample design or procedures.

The census/survey choices

Assuming that some pattern of direct census and survey enquiries is needed in addition to the information that can be drawn from administrative sources, the government statistician has to choose options carefully in the light of the characteristics of the different types of enquiry that could be undertaken.

To illustrate, in terms of the standard enquiries of households, farms and of business enterprises and their establishments, these options might be characterised as follows:

Censuses (including sample inquiries) of the detailed benchmark type — eg as with a population census or the major decennial censuses of the *World Programme of Industrial Statistics*. These

benchmark censuses establish the framework

- are *major undertakings* providing detail at expense of timeliness;

- *create a framework* for more frequent enquiries;

- yield fine data item *detail* (in the case of business enquiries in respect of the most detailed statistical unit level they are a good basis for detailed analyses and studies, particularly if supplemented by special investigations)

- in the case of business enquiries, may be designed to provide information relating both to individual physical *establishments and the enterprise as a whole*;

- may be *suitable for regional studies* — especially for data on physical units such as households or business establishments, published in geographical detail, or if the individual records or the small area 'building blocks' can be manipulated to provide information for user-defined areas;

- *will not always give comparability over time* because of changes in design, classification systems etc when censuses are conducted at long intervals. They are oriented towards *cross sectional* studies rather than to comparisons over time. As such they are likely to be tremendous mines of information. But their value can be greatly increased when they are repeated at intervals, with attention to the comparability problem so as to enable *structural change* to be assessed.

Annual Enquiries

intercensal continuity

- may be censuses or sample enquiries

- should provide linkage between successive periodical benchmark type censuses if key items from the detailed censuses are also collected in intermediate enquiries;

- level of detail may be limited;

- time lag in production of tabulations may be less than for benchmark census, but is likely to be relatively lengthy;

- scope and coverage might be less ambitious than in a less frequent benchmark inquiry;

- annual enquiries may result in more effective benchmark enquiries, because of the continuity of reporting, maintenance of trained statistical staff and enumerators, availability of an up to date population list (eg register of employers), proven procedures etc;

- on the other hand an annual enquiry may tie up resources fully and continuously and additional resources for a benchmark enquiry may be difficult to obtain.

Annual Or More Frequent Current Indicator Enquiries

current indicators

- higher level statistical units (eg the whole business instead of each business location) may be necessary to ensure speed of response and comprehensive, economical coverage;

- the data might be used for carrying forward census data but only at broader levels of classification;

- it is important to provide for consistency of reporting for different periods in a coordinated system of benchmark and annual, quarterly and monthly inquiries, but absolute comparability may not be possible if different units are employed;

- sometimes only selected data are collected in such frequent inquiries and are used (with care) to extrapolate other benchmark data.

Relationships between periodical census and annual or more frequent enquiries

Some countries have infrequent (eg decennial or 5 yearly) censuses and have subsequently complemented these with annual or more frequent enquiries. This is almost universally the case with population censuses and surveys.

Others have long standing annual enquiries and are developing more comprehensive and detailed censuses at less frequent intervals. Some countries have a rolling program of periodical censuses for some industries with annual and more frequent enquiries for others.

Opinions vary as to whether the emphasis should be placed on the decennial benchmark enquiry or on the annual enquiry. Should the annual enquiry be regarded as the CENTRAL enquiry of a coordinated system for the collection of data and a benchmark enquiry be regarded as a supplement to the annual enquiry? Or vice versa?

Probably the answer to this question for a particular country will depend on the way it has gone about the job in the past, as well as on user's needs and on resource considerations.

In countries where, to conserve statistical resources, the annual enquiries are limited in scope and coverage or are sample enquiries, it is also natural to treat them as subsidiary to the benchmark enquiries.

In any case, the benchmark enquiries can be expected to perform functions which distinguish them from the more frequent enquiries:

(1) to establish a detailed profile of the structure of the industry or sector;

(2) to provide a broader and more accurate array of information on activity than would normally be feasible on an annual basis.

Given this detail, the benchmark enquiry can supply the framework needed for undertaking special surveys and can provide the basic information required for the evaluation and adjustment of current indicators.

In those countries where annual enquiries have not yet been instituted, an enquiry of the type envisaged can provide an appropriate basis for their organisation and development.

Preliminary results and rolling revisions

Another measure needed to deal with the time lag problem in continuing systems is to provide for a succession of estimates in both the direct collection programmes and in the subsequent analytically compiled results. For a survey this might take the following form:

(1) *preliminary results* (estimates with minimal detail prepared before all responses have been received or processed;

(2) *final results* (prepared after all responses have been processed and the compilations have been *closed off*).

Sometimes, in a continuing regular survey, there may be still further revisions for previous periods incorporated in the time series when the results for the current period are published.

revision cycles

This may be irritating for users, particularly if they are not made aware of this possibility. **But it is important that producers of statistics should not be inhibited in aiming to produce both timely advance figures and more accurate later estimates incorporating the best information available at each point in time.**

Of course, this should be done in such a way as will give the user the clearest possible picture of the rate and direction of the change, allowing for any evident tendencies for particular series to be revised upwards or downwards in successive revisions.

The likelihood of *rolling revisions* to be made over a number of periods increases as *indicator series* are nested between large detailed periodical *benchmark* enquiries such as industrial censuses. Thus the benchmark final figures may be *extrapolated* forwards, quarter by quarter, for some years until the next census replaces the latest survey-based figures. The whole extrapolated series may then require adjustment up or down to link the previous benchmark with the latest benchmark and a new extrapolation sequence begins for periods ahead.

To sum up

A well-developed national statistical system will have a continuing nested sequence of censuses and surveys providing a succession of prompt summary information and later finalised more reliable and detailed results. In general, the paramount consideration in the development of such continuing systems will be that they provide reliable measures of change over time.

In later chapters we will see how the system becomes very much more complex with the analytically compiled data systems, such as the national accounts, which involve rolling revisions to series which are themselves a compound of different series each of which have to be reconciled with one another at the same time as successive revisions in the separate series are incorporated. Given the many directions from which new data may enter into such systems to displace or modify the picture built-up from earlier but less reliable sources, our NSS looks to be quite a complicated animal. At this point it is sufficient to note that the statistical representation of the nature and state of an economy with respect to any period or moment of time is only gradually resolved over the five or ten years subsequent to that time.

CHAPTER 7

INTEGRATING ENQUIRIES ACROSS SUBJECT MATTER FIELDS

Still concentrating on the statistical production process, this chapter elaborates the picture of continuing census and survey sequences to show how a comprehensive national statistical system will call for the processes of information production to be integrated or co-ordinated across a wide range of subject matter. How does this process of operational integration come about? What are its elements? How integrated can you get?

How does it happen?

INTEGRATION (of statistical systems) involves making diverse elements into a functional whole.

INTEGRITY (of statistical systems and statisticians) denotes wholeness, soundness of function, honesty and candour.

An INTEGRATED STATISTICAL INFORMATION SYSTEM should have these qualities. An NSS is poorly integrated if it is wanting in these qualities.

It is worth saying at the outset that the integration of statistical systems, be it the manufacturing statistics programme with the distribution statistics programme, or the police statistics programme with the courts and corrective institution statistics programmes, comes about as the **natural pressure from users of statistics to ensure that information systems about the different actors and activities in society can be related, just as they are related in the real world through social interchange**.

Government statisticians respond to this in the best way they can, given the institutional environment in which they work. We have seen however, that there can be a natural resistance to pursuing the objective of integration because every change in the specifications of a data collection (albeit to make it consistent with some other) is likely to involve breaks in time series and a less positive representation of the changes we are concerned to measure.

When this professional argument against making changes to data collection systems is coupled with institutional inertias, progress towards more unified statistical programmes is likely to happen only as government statisticians *catch a vision* of the possibilities of a national statistical system and give a high priority to the pursuit of integration.

What are the elements of integration?

Chapter 6 indicated how a continuing data collection system may evolve to cover a given population of statistical units such as establishments classified to secondary industry.

In the terminology of the *data box* (page 26 of Part 1) we have concentrated on the TIME dimension of a national statistical system's processes.

We can now elaborate this picture to consider how the system might be made more comprehensive with respect to the SCOPE and COVERAGE of the STATISTICAL UNITS covered (eg to embrace the full range of industries or household types) and with respect to the CONTENT of DATA ITEMS (eg to cover borrowing and lending as well as production activities).

The various parts of a whole can be said to be well *integrated* when they are related to one another in a consistent and complementary way.

We appreciate the superb level of integration in our high performance car. We may be far less impressed with the integration of the metropolitan road system where our car has to operate along with trucks, buses and bicycles. We are judging the system by the way the sub-systems and related systems complement one another as a whole.

What would we be looking for in a national economic census and survey system?

commonality of statistical unit and data items — allows ready aggregation and disaggregation

Statistical reporting systems might be said to be integrated with each other if the statistical collections are consistent and complementary, with no overlapping or gaps in coverage, (the STATISTICAL UNIT dimension) and if they produce a range of data defined according to a common system of concepts (the DATA ITEM dimension) and with units and data items classified in a consistent or related way (the CLASSIFICATION dimension) so that they can be aggregated to form meaningful totals for the society or category as a whole or for parts of it.

We would be looking also for the aggregations to be done at regular points of time, or in respect of comparable periods, ie we would look for consistency in the timing of collection and processing operations (the TIME dimension).

How integrated can you get?

There are degrees in which requirements of internal consistency might be imposed on a complex of statistical reporting and compilation programmes. A fully integrated system implies a unitary design from the ground up, or at least the utilisation of common elements (such as common definitions of reporting units and a range of common data items) in the separate programmes.

A statistical coordination agency may achieve coordination in the reporting programmes of independent statistical agencies so that the agencies work consistently without duplication and do not overburden the statistical collection respondents.

co-ordination may not ensure integration

But the total statistical reporting system which emerges within the constraints of its controls may be far from being fully integrated. Each individual programme may meet its standards, but between them there may be differences in scope and coverage and there may be incompatibility for such reasons as the fact that they are not based on a common list of statistical units, consistently defined.

For a variety of reasons, a system of separate enquiries may not permit ready aggregation across different statistical enquiries for the same geographical area or category of statistical unit. Comparability may be possible only at aggregate level (eg) for national totals (after special adjustments have been made) but not for small geographical areas or individual units.

Where the establishments or companies belonging to complex corporate groups are the subject of separate surveys, there is a risk that data obtained for establishment units will not be consistent with data reported for related units at another accounting level.

This can be very difficult to check if different surveys are made of these accounting records by different agencies at different times. Thus, adequate safeguards are essential to avoid gaps and duplication when enterprise and establishment data are collected and associated in the analysis of business activity. The nature of this problem will be examined more closely in Part 3.

How widely can we integrate?

While it might seem a reasonable goal to have an integrated statistical information system to cover all production activity, for example, (as would be required if the domestic production account were to be built up directly from establishment data), even countries with highly developed government statistical systems are a long way from achieving this.

Excepting small uncomplicated countries, comprehensive coverage is commonly achieved only in terms of the enterprise unit (which can be too broad a unit to be satisfactory for detailed analysis of economic activity by class of activity and by geographical area).

The coverage of establishments may be fairly complete across primary and secondary industry but coverage of services is likely to be limited at best to wholesaling, retailing and some of the associated services. Similar examples of such problems can be found in all fields of statistics.

a 'clean slate'?

There is then the further problem that many statistical programmes have histories and conceptual bases going back beyond the period when the demand for comprehensive planning data had much influence.

In a sense, many of the more recently developing countries have had a much less difficult task in integrating their statistics programmes because they were able to begin this at the design stage as they commenced collections on the basis of international standards. Other countries have been faced with the problem of making major changes to unify separate programmes, each of which in themselves have been satisfactorily meeting their original goals.

Accordingly, in long-established statistical services, integration may be best approached on a step-by-step basis, concentrating first on economic statistics relating to combinations of industrial sectors where the uncertainties arising from lack of integration are causing the greatest difficulty for statistical users. Integration of social statistics has not proceeded as fast or as systematically for reasons that are discussed in Part 4.

integrating economic censuses and surveys

Thus, a first priority for many countries might be given to integrating collections from mining, manufacturing, wholesaling and retailing establishments — these activities tend to be closely associated (eg they are common combinations in multi-establishment enterprises) and there are difficult problems of reporting system coordination arising from that. There is also a frequent demand for the association of statistics in respect of them (eg someone analysing the location of manufacturing industry is likely to want fully comparable data about selling outlets).

As a further step, construction statistics might be integrated because of the problems of drawing a clean line between manufacturing and construction. Agriculture, on the other hand, might be a matter of less concern because agricultural units tend to stand more clearly on their own in most countries where they are almost all single-establishment enterprises and (eg) mixed agricultural/manufacturing enterprises are much less common than mixed manufacturing/ wholesaling enterprises.

incorporating administrative data systems

The boundary of a project to integrate enterprise and establishment statistics is likely in the first instance to be limited to the **direct** statistical reporting systems. It may be very important (eg for the development of national accounts) to bring administrative by-product data systems within the scope of an integration programme. But the inclination of administrative agencies (eg a Taxation Department) may well be to regard their programme

management requirements as overriding the external requirements of central statistical offices and national planners.

Consequently, while co-operative arrangements often can be mutually beneficial, it may be necessary to accept the data content of some agency records as given. It may also be necessary to work with very limited access to the records and to the data processing operations performed on them in the administrative agency. Perhaps only aggregated data is provided, or if individual data are provided it may be with identification removed.

Such lack of access to administrative records is likely to be a problem with income tax data in particular. Thus it may be impossible to classify income tax returns to industry positively on a basis comparable with their treatment in other enterprise collections. This can introduce uncertainty in reconciling income tax data for enterprises with economic census data tabulated for enterprises. But in some countries at least, it has been possible to use taxation data as a substitute source of quantitative data for particular units (eg small single-establishment enterprises) omitting them from economic census collections.

Theoretical frameworks

the 'blue book'
— a master framework

The various agencies which have developed statistical collections on different bases for their own special purposes may have different views as to the desirable uniform standards to be employed. What is needed is a master statistical framework which will help resolve consistency issues by providing fully developed descriptions of statistical units, data items and presentations. The *UN System of National Accounts* (SNA) provides such a general purpose system for economic statistics, which anticipates a wide range of possible information requirements. The necessity of having to bring the data in its accounts into balance imposes a consistency check on the component compilations.

It may be that alternative theoretical frameworks (and, certainly, elaborations of the SNA framework) will need to be developed as a basis for designing versatile integrated enterprise/establishment data systems. But the theoretical basis of the SNA is accepted enough for integration programs to proceed confidently in respect of business units.

There is no counterpart internationally accepted framework for non-economic social statistics, although some studies such as Towards a System of Social and Demographic Statistics (SSDS) at least bring out the connections between the various social and demographic fields and the cohesive money-value accounting of SNA-based economic statistics.

Standardisation of concepts

On the basis of a general theoretical framework, such as the SNA, standard definitions and classifications are defined for statistical units and data items.

While it is reasonable to expect that there will be a call for data items to be collected in some collections and not in others, integration would require that some items be collected in all collections, using uniform definitions.

In the case of economic statistics, for example, key items that would enter into a reconciliation between an enterprise return and returns for component establishments would be shipments (sales and transfers), stocks, purchases and selected expenses (items involved in calculating value added), employment, salaries and wages and fixed capital expenditure. These are items which would require a common conceptual basis in collections from establishments in respect of different industries and from enterprises.

Other items, consistently defined, might be added to this list in the case of enterprise collections; for example, gross operating surplus (which, with the addition of salaries and wages and net indirect taxes, gives contribution to gross domestic product) and other financial items, such as interest receipts and payments, depreciation, and expenses not available at the level of the individual establishment.

With such information, some countries have (with some difficulty) linked income data from taxation sources with production data from the establishment collections.

Vertical and horizontal integration

Integration of statistical information can have two dimensions, vertical and horizontal. The distinction is a familiar one with respect to business enterprises. Thus the activities of statistical units can be integrated horizontally across the different kinds of production activity (such as mining, manufacturing, wholesaling etc) and also vertically in terms of the different levels of business organisation.

Analogous distinctions can be drawn in regard to statistics for which the unit can be the individual, the nuclear family, the household or some broader social group, all of which can be surveyed partially in different enquiries.

reassembling data at different accounting levels

Integration objectives in this context should be to provide for the data collected in respect of individuals or establishments to be reassembled at an aggregate level as data relating to the broader units of which they are part.

The user should also be able to call for cross-tabulations, eg tables in which data are given for establishments classified by industry showing separately those which are owned by a) enterprises classified to the same industry, and b) enterprises classified to other industries. Such transformations between alternative systems of classification and aggregation could be very useful for

particular purposes of analysis and the integrated system should provide for them if the information collected is to be fully used.

Integrated collections

The most complete and positive integration of data, both vertically and horizontally, is likely to obtain when data are collected from each level in a hierarchy of statistical units in a single collection process. For example, in industry statistics, this could require each enterprise to supply data for the enterprise as a whole, reconciled with the total of the data provided for each of its component establishments.

Collections which are made at one time from some units of an enterprise and at another time from the remaining units classified to other industries, and collections which obtain totals for higher level units indirectly by addition of lower level units data, are open to inconsistent reporting, gaps and duplications. Nevertheless, accepting some uncertainties arising from partial approaches at different times, it is possible to have an integrated system if all units are positively identified and coverage is complete.

need complete list of units

Thus, much is gained if there is at least a full canvass of both higher level and lower level units in the course of preparing and maintaining a central register and if the collections by various agencies are based on this common list so that the collected data can be aggregated without gap or duplication.

Instruments for integration

It will be apparent that successful integration of statistical collection will hinge on the availability and general adoption of:

Conceptual system elements

conceptual integration

- a *comprehensive conceptual framework* for the statistics, defining and relating the key concepts of statistical units and data items;

- *standard definitions of statistical units and data items* across the integrated collections;

- *standard classifications* applied to these units.

Operational elements

operational integration

- a *central register* of statistical units in which all units are classified consistently in accordance with the standard classifications;

- *common operational standards* relating to quality, timing etc, which may be embodied in integrated computerised processing systems.

The central register, recording hierarchic links between different levels of units, is an essential basis for the elimination of overlapping and unintentional gaps between different data collections. It also ensures that each individual unit is classified in the same way in each compilation and opens up further possibilities for establishing various links between different sets of statistical data from different sources, which can greatly increase the amount of useful information that can be derived from the total stock of data available.

Wider still?

We have illustrated the integration process with emphasis on the problems and issues arising in the integration of economic censuses and surveys.

The pursuit of integration in forms appropriate to the different subject-matter fields of national statistical systems is a basic concern of government statisticians. A national statistical system is more or less systematic to the extent that the designers and developers of data collection and processing programmes succeed in this.

But to put this into perspective, a thoroughgoing integration of even the basic economic censuses and surveys has been successfully completed by very few countries. For the rest of economic statistics about society the statistical operations are likely to consist of a great variety of programmes which are only loosely meshed as mutually supporting elements of a comprehensive national statistical system. And, as subsequent chapters will show, there is much more to statistical integration than just the systems for producing a well-integrated output.

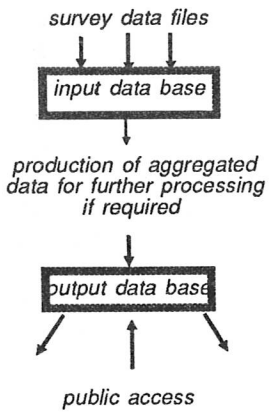
CHAPTER 8

MANAGING THE STOCK OF DATA — THE DATA BASE APPROACH

In Part 2 of our 'guided tour' we have so far been given a progressively more elaborate picture of how, in an NSS, files of current and historical data should be produced and maintained on the basis of common conceptual and operational standards. We now consider how the concept of integrated statistical file systems and the concept of data bases takes us a further step towards the objective of providing a better integrated statistical information service.

We have been discussing in Chapter 7 how we might systematise statistical reporting and compilation programmes so that we obtain the information we need, defined and classified in a consistent and complementary way over a wide range of subject matter. We have pictured the NSS assembling a large number of data files containing the 'clean' edited data from the statistical units covered in each enquiry and/or the aggregated data subsequently assembled for publication. Integration has been provided for to the extent that the collected and compiled data meet the specifications of our statistical framework. But we have stopped short of considering the operational questions of how we might hold such a system of statistical files so that users might obtain maximum benefit from them. This is an important part of the general problem of data management which has led to a data base approach to statistical systems. But first we need to define some of the terms used in this connection.

Elements of the data base approach



Definitions vary, but the following suffice for our purpose:

Data file — a set of consistently defined observations (relating to a single subject or with at least one characteristic in common) recorded on a common medium.

Data base — a set of related data files.

Thus, in terms of our discussion to date, by *data base* we have in mind a system of statistical files relating, (eg) to a broad class of statistical units.

One example of a well integrated data base is the data maintained by the Australian 'Integrated Economic Statistical Information System' (IESIS) which has been used since 1978 for the full range of production censuses and surveys. This is essentially an input data base. It may be linked to an output data base, holding data available for public access from all of the Bureau's economic and social statistics programmes. But needs and possibilities change and IESIS will give way to something better!

In a broad logical sense the total holdings of data on households in an NSS could also be described as a data base, even if the files may be poorly related operationally.

More restrictive computer systems' definitions would prescribe the nature of the relationship between the data in a computerised data base as necessarily having a common definition and description and being structured in a particular manner. The term *data base management system* then is used generally to apply to a system which

provides for the definition, storage, retrieval and maintenance of data in computerised data bases, ie those which necessarily require very explicit and detailed specifications and force close attention to issues of relatability which are beyond the practical limits of manual systems.

Our discussion needs to leave open the question of how well data files need to be related to constitute a *data base*. In fact it is perhaps better to speak of a *data base approach* of applying *data base technology*, to the extent that it is possible and appropriate, to the complex of programmes and projects encompassed by a National Statistical System.

so what is data base technology?

Data base technology and data base management systems concern the use of computer techniques for storing, accessing and processing data more efficiently once it has been acquired.

The application of data base technology can be very complex or very simple depending on the scope of the data base information holding and on the use to which the system is put.

For example, the concept of an electricity accounts system or even an airline booking system or an inventory control system is reasonably straightforward in the sense that it usually deals with a finite set of items and finite set of data. The inputs and outputs of the system are rigidly defined and the actions that are carried out on the data are also well defined.

The bounds of a computerised statistical data base should be confined and clearly defined but the bounds of a statistical system are not always clearly defined and the outputs which will be required are inherently difficult to define firmly as the supply of information to users tends to raise new requirements.

Certainly one aim of integration is to provide statistics that are more comparable and more meaningful across a wider range of statistical series so that further processing, manipulation or modelling can be carried out. Publication of predetermined tabulations is then not the sole aim of the production system.

The system will be called upon to produce new tabulations, cross classifications and arrays of the data it contains. It must also even 'relate' inconsistent data, so that users can at least access all that is (potentially) relevant (and then wrestle with reconciling it or selecting from it).

Sensibly, integrated statistical file systems should be geared to improved methods of data outputs and it is in this area that data bases connected by common data index/search/catalogue facilities can provide a better service.

Without pursuing the possibilities of applying computer technology further here, it is worth considering what in essence the data base approach involves for statistical production. This requires the introduction of some further terms used in this connection. Thus a distinction between three *levels* of data is important:

levels of data

Raw data (as gathered in a survey)

Clean (micro) data (the unit record data after coding and editing)

Aggregated (macro) data (these may be assembled from the micro data or may be direct aggregate estimates not always fully consistent with related micro data sets. As we shall see in Part 3, this is commonly the situation in practice in the case of many national accounting aggregates).

Another element of the data base is *meta data*, such as descriptions and definitions of the data, the survey frame, classification coding tables, editing and aggregation rules and table frames — in other words, information about the data. This might best be maintained in a *data directory/dictionary,* a readily accessible computer file to assist in servicing information requests and to point to the location of data in the data base.

The construction of such a facility helps to reveal inconsistencies (and helps in monitoring them also), provided the different operational measures of each underlying concept are linked.

What goes into a statistical data base?

As processing proceeds through the stages of converting the raw data into acceptable micro data and then into aggregates for dissemination or access, there will be an interaction between the survey data collection and processing systems on the one hand and the statistical file system on the other.

As with a repetitive enquiry (see Chapter 6) the statistical file system may contain data from previous surveys in its data base. In the case of more extensive systems it may contain data from a number of different surveys and other sources of data relating to the statistical units in the data base. A central register of businesses may be regarded as logically a subset of a data base of business unit census and survey information.

From the users' perspective

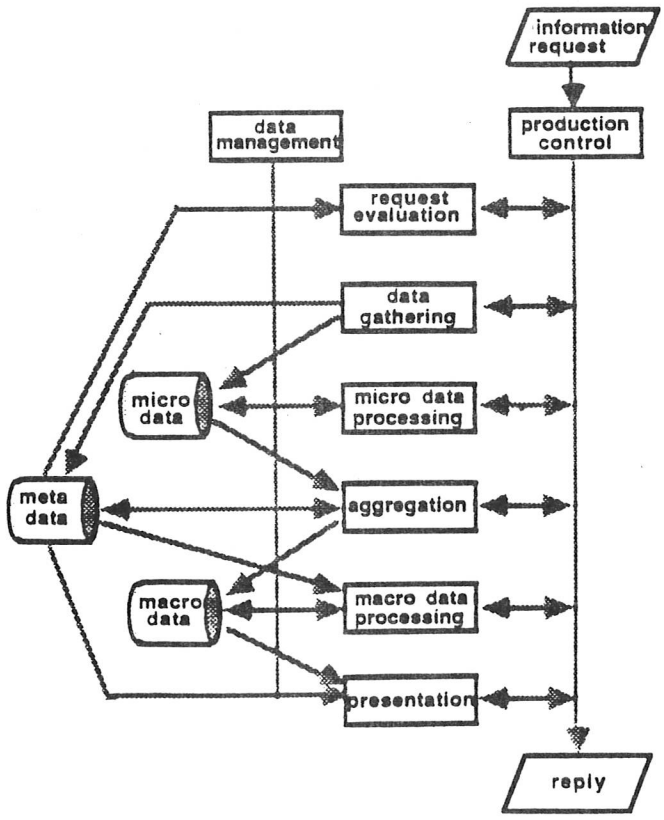
The data base approach is essentially one of maintaining as much related data as possible in a readily accessible form. **The approach recognises the fact that statistical data are information *capital* that can be used over and over again to produce additional information**

needed by users and that it is important to facilitate this. Bringing together data collected in different contexts may also enlarge the information stock by synergetic processes — where the sum of the information product is greater than the sum of the separate parts.

It is not enough to produce the prescribed tabulations of statistical aggregates. Micro data need to be available so that new macro data can be generated in response to continuing demand arising from new users and usages of statistical data. Not all these users are known and not all the usages can be fully anticipated at the time of data gathering. As we have noted, **the use of statistical data in decision making and research more often creates new demand as the initial requirements are met.**

How is the statistical production process affected?

The main operational effect of the data base approach to statistical production has been to separate the traditional statistical production routines from the management of the data base of micro, macro and meta data. Thus, as depicted in the *data oriented model* (Veim and Sundgren, *Bulletin of the International Statistical Institute* Vol 48, 1979, part 3, p92) data management for a whole range of surveys may now be handled by a data management unit responsible for maintaining data definitions, structures and linkage information, as well as the storage and retrieval of the data. Note that *meta data,* data about the data, is a key element.



There are alternative ways in which the data bases may be physically arranged. Thus it may prove more economical and convenient to separate an output data base containing only non-confidential macro or meta data in a form readily accessible to external users. Processing routines will generally use the (micro) input data base as input to the aggregation routines which generate, up date or modify the public-use data base. The data in the output data base may also be manipulated in various ways to meet user requirements.

While the data base approach is being pursued in most countries which have the technical facilities and capability, there are philosophical and technical constraints relating particularly to confidentiality, security and access rights. While not unique to data base systems they do cause operating procedures to become quite complex.

Confidentiality

While some non-sensitive kinds of information are excepted from the non-disclosure provisions of some countries' legislation, most unit record data collected and processed are, by their very nature, confidential. The use of data base technology increases the confidentiality problem because more data items or classifications of data items can be made available readily to the user in respect of tailor-made classes of statistical units. A statistical organisation therefore has to work out very carefully its rules as to what level of aggregation its data can be supplied.

Access rights and security

A supplier of data working in a data base environment is also faced with the problem of deciding who should be allowed access to which data under what circumstances. The data base system using the computer has to be able to differentiate between the users by means of a series of passwords for data access rights which ensure that users can only obtain agreed sets of data or can only use it for authorised purposes — eg some can both *amend* and *interrogate* data, others can only interrogate it.

In any large data accessing system there must be a general set of rules concerning the security of data, how it is stored and backed up. The more data are available for use, then the more comprehensive has to be the system so that in the case of accident, data can be reproduced or replaced and the data service is not destroyed.

To Sum Up — *(in the words of Veim and Sundgren — op cit)*

"The data base approach centralises the management of data into one specified function. The data becomes a common asset, shared by multiple processes within the survey production process.. It is "natural" to extend the idea of data sharing from the processes of one survey to the processes of several surveys, either repeated surveys or data related surveys.

This extended sharing concept will lead to a new production structure in which the data and "its" survey participate in a many to many relationship between a set of surveys and a (set of) data base(s). In this way the data base approach for statistical data processing could promote survey integration and recycling of data. A side-effect will be pressure for better coordination of surveys and more available data definitions.

From the above, we can expect that the data base approach to statistical data management will stimulate:

- *centralised control of data and data descriptions;*
- *integration of data and meta data;*
- *standardisation of functions and interfaces in statistical production systems;*
- *integration of data from different surveys and time periods; and*
- *recycling of data*

This can lead to :

- *more flexible and robust systems which will be easier to develop and maintain;*
- *simpler and more rapid data access;*
- *less unnecessary collection and storage of redundant data; and*
- *more consistent data that are easier to combine over time and subject.*

Data base techniques seem to offer significant contributions to the overall goals of a statistical agency:

- *better service to users of statistics;*
- *less burden on the repondents; and*
- *more efficient production processes."*

We would add to this summary that data base technology demands a considerable sharpening up of data management principles and practices relating to confidentiality, access rights and security.

CHAPTER 9

USING THE DATA STOCK IN ANALYSIS, SYNTHESIS, PROBLEM SOLVING AND SYSTEM FEEDBACK

To this point we have pictured the processes involved in collecting a range of basic data sets, in both macro and micro (unit record) data form, and holding and managing them in data bases, ready to be used.

In this chapter we consider how further higher-level statistics are derived (and may be incorporated in the data bases) by processes of analysis or synthesis. Finally, we consider how the statistics are applied by users for research, policy advising and decision-making and how this experience may be applied to the further development of the information system.

Basic statistics

We have been discussing statistics derived from the aggregation of data relating to individual statistical units — businesses, households, government and other institutions. Such series are commonly termed *basic* or *primary* statistics. They are produced in the NSS by processes which involve:

(1) data gathering (from direct surveys or as an administrative by-product);

(2) coding and editing;

(3) aggregation and some analysis; and

(4) presentation of statistics (commonly with entry into a data base after step 2).

can be complex products

Characteristically this involves massive data collection and processing operations in which micro data are archived at the end of an initial production process while macro data are made available as the *ready to use* end-product.

Notwithstanding these characteristics, so-called *basic statistics* are often complex products in which professional statistical skills are called for, including those of analysis and interpretation (see *analysis in the production of basic statistics* below).

Another qualification to note is that the distinction between basic statistics from *direct surveys* and basic statistics from *administrative by-product* sources is not always clear cut. The characteristic difference here is that the Statistician can control the specifications of the direct collections while generally having to take the

specifications of the administrative by-product data as given. Sometimes access may only be to aggregated data from the latter source.

direct survey and administrative information may be inter-dependent

However, the differences tend to be matters of degree. On the one hand, the Statistician has to work with the accepted accounting and other conventions of respondents and with their general reporting capabilities while, on the other hand, the Statistician may well influence the definitional and processing standards of government agencies, so that the administrative information systems of individual agencies are satisfactorily aligned with the general purpose statistical information systems of the NSS. In any case, many direct surveys, particularly business surveys, will depend heavily on administrative sources for framework purposes, particularly to identify *births* of statistical units.

Higher level derived or analytical series

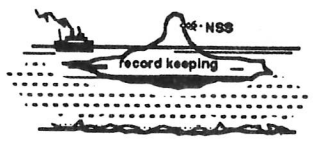
Having already somewhat muddied our image of the NSS data bases consisting of directly observed and aggregated data about individual statistical units, we need to appreciate that statistical series of a substantially different character will also be held in NSS data bases. These are the series (such as national accounting estimates and price indexes) which are created by indirect methods.

At the frontier of statistical development these may be little more than *back of the envelope* estimates. Generally they use relatively small volumes of data but are likely to involve complex assumptions and the manipulation of different statistical series in association. These days much of the work is done on micro-computers with access to the data bases of both directly collected and derived statistics.

NSS — the tip of the iceberg

Our picture of the total National Statistical System presided over by a National Statistician now takes on the rather complex form of a hierarchy of systems which are to varying degrees interdependent, tied together by a large number of interconnected processes.

This includes a growing number of national accounting and other derived official statistical series in which the government statisticians are themselves users of statistics which they must analyse and interpret so as to take action to solve a problem associated with the production of statistical information of a higher order.



An essential part of the hierarchy of the NSS system will of course be the information systems of the businesses, government and other institutions from which the basic data are drawn. Like the base of the iceberg, they may not be visibly part of the NSS but they keep it afloat!

statistics as a stock of material for producing further statistcs — used and re-used, but not used up

The *iceberg* is anything but inert, however, as all the systems are interacting. All the statistics, including analytically produced series, may be inputs to statistical processes aimed at producing answers to the general and particular questions of users. **With the constant turnover of the statistical stock in this process there is, in a sense, no statistical end-product and certainly the end is not at the point of the issue of a publication by the statistical office.**

Even complex statistics such as those included in national accounts series will undergo further transformations by crude or sophisticated methods in the hands of users. As with other statistics, users generally only want them in order to project ahead from them, because their planning decisions must relate to future periods.

What is analysis and interpretation?

analysis — take apart

synthesis — put together

Some discussion of these terms now seems appropriate as we are straying across the line where the functions of a government statistical service might be considered to end and the functions of researchers, planners, advisers and decision-makers might be said to begin. In particular, it has been held by some that *analysis* is properly within the boundary of the NSS but *interpretation* is not.

Strictly, to analyse information is to take it apart and look at the bits, as opposed to *synthesis,* which is to put a lot of bits of information together to create new information out of them. In this sense of the word, analysis in relation to statistical processes involves:

(a) *examining the components of an aggregate series* in order to check consistency and validity (eg was the movement a general one or was it due to some exceptional case which, on checking, proves to be incorrect).

(b) *understanding the meaning of movements in statistical aggregates* in the light of associated events (eg to infer that a rise in Gross Domestic Product would have been much greater if farm product had not been depressed by drought conditions).

The first type of operation is fairly standard in the output editing processes of statistical offices and sometimes some analysis of this kind is incorporated in statistical press releases where a government statistician is

prepared to offer explanations of movements in aggregate series by making some reference to movements in components of the aggregates or of other related series.

This is purely descriptive analysis. The kind of analysis suggested in (b) may involve interpretation of a widely accepted and conventionalised nature, as when series are adjusted to remove seasonal and other elements in order to perceive trends. Alternatively, they may involve expressing subjective judgments as to cause and effect.

This latter is much more hazardous. Statisticians usually avoid making comments of this kind in their publications unless they are based on factual evidence. In particular they will avoid comment on government policies and their effectiveness.

Synthesis in relation to statistical processes is usually also labelled as *analysis*. It would be illustrated by:

(a) *summarisation* of data so as to provide generalisation of higher order from particular series (eg to bring together, in a summary statement or graphic display, information published in different but compatible collections);

(b) *production of new statistics* by making use of the known relationships between available data series and the series to be estimated.

As opposed to (a), which is the more or less mechanical addition of data by the normal compiling process, (b) involves indirect and relatively complex analytical methods. Thus we might refer to the national accounts and indexes of industrial production as *analytically-produced* statistics or as *synthetic* series.

In fact, many national accounts series do involve breaking down a national aggregate into components i.e. analysis rather than synthesis. As we examine the origins of the body of economic statistics of the NSS, concrete illustrations will be provided. Essentially such statistical production involves inference (eg as to relationships between series, as to which conflicting indicator to believe etc.).

Interpretation is the step in the statistical process in which inferences are drawn (where conclusions are reached about the meaning of the figures in relation to a problem to be solved). But, as suggested above, this goes hand in hand with analysis — both in pulling statistics to pieces in order to draw out their meaning and in putting statistics together in the process of producing new statistics by analytical means.

Attitudes to analysis and interpretation

The view, that *analysis* is legitimate for a statistical service but *interpretation* is not, is deeply rooted in the departmental philosophy of some statistical agencies. **This view holds that factual description and interpretative judgments should be kept separate and that interpretative advice is not the responsibility of a national statistical agency; rather its function is to supply information (including statistical analysis) from which users may draw inferences.**

In this way, the credibility of the NSS as an independent impartial service can be preserved. Too close an involvement in government policy issues may make the statistics less believable and therefore less useful.

For example it has been said of George Jaszi (former Director of the US Bureau of Economic Analysis) that 'he views the excellence of BEA's work as the shield behind which politically sensitive statistics, such as GNP and the leading indicators, can be prepared without partisan interference. Further, he has viewed the provision of direct policy advice as the sure way to invite such interference' (Survey of Current Business, February 1985).

On the other hand, when statistical effort is closely associated with the programme missions of Departments with policy responsibilities, government statisticians will be more concerned with trimming their sails according to their technical competence and resources capability at each point in time than with the question of what it is *proper* for an official statistician to do.

Such agencies will be firmly led to develop their service capability in areas of both analysis and interpretation in which they are able to contribute their intimate knowledge of the data. They consider that analysing and interpreting the figures must be carried out in close touch with the roles of collecting and validating figures and, perhaps of developing the techniques of analysis. They will regard themselves as part of a policy advising team jointly responsible for the facts and inferences on which are based the decisions, policies or plans which are the objective of the whole statistical process.

different approaches — same goal

These differences in points of view generally reflect differences in the NSS's institutional environment, particularly as to the degree of centralisation/ decentralisation of the system. **Certainly the ultimate aim of a national statistical system, whether centralised or decentralised, is to help solve problems through the analysis and interpretation of whatever statistical material can be assembled.**

Analysis in the production of basic statistics

In any case, statisticians cannot avoid analysis and interpretation of data even if they specialise in direct data collections.

In the first place it is a necessary phase of the statistical production process to examine the output in a systematic way in order to check its credibility, both in regard to its internal consistency and in relation to other external evidence.

In the second place it is necessary to undertake regular analysis of continuing surveys in order to modify and revise them as weaknesses are revealed.

was the 'sample-take ' representative? were the questions answered correctly?

So a properly developed evaluation programme will check regularly on *sampling errors* (through the regular computation of sampling errors), and *non-sampling errors* (eg through comparisons with independent estimates or through re-enumeration).

One source of error, which is particularly important in surveys for which early provisional estimates are required, is the *non-response* error which arises from having to impute (ascribe) estimates for returns which are not received in time. This can arise when the returns received at the cutoff point are not representative of the total sample.

Computers can assess the increase in sampling error that early estimates are subject to in comparison with final estimates. Analysis can also identify units for which, if they do not respond, the imputations would have an unduly large impact on the final estimates. (These may either be large units whose statistics have a major impact on the current estimates, or medium sized units for which experience indicates that it is particularly difficult to impute with reasonable accuracy because of the volatility of their activities.)

Regular analyses can also identify, for special attention, units that have been non-respondents or late respondents in two or three consecutive surveys. Continuing efforts to focus attention on such weak points in a collection can do much to improve the speed and reliability of a survey.

Analysis and interpretation to produce synthetic series

We have argued that analysis and interpretation of data are inseparable from the operation of producing acceptable basic economic statistics. What has been described is a process of obtaining supplementary information which is needed for judgments to be made as to whether the results are valid enough for publication and as to what improvements are needed in the system. This is internal information (although some of it will probably be published as an aid to the user).

seasonal adjustment sharpens interpretation

But analytical processes are also used in producing estimates for publication. Thus many statistical agencies produce regular statistics by routine analytical techniques which are based on the characteristics of particular series

(eg to produce seasonally adjusted series) or on a knowledge of the relationships between available data series and the series to be estimated. Many of the national accounts series, price indexes, demographic and econometric forecasts are regularly produced in this way.

The methods employed are characteristically indirect and complex by comparison with the direct processing of large-scale files in economic censuses, surveys and administrative by-product collections. Typically they require elaborate rather than massive data computations, working with summary data rather than individual record data.

As our later examination of the NSS as a body of data will illustrate, the methods involve the adaptation and manipulation of various data series together, where no one of the data series on its own adequately represents the stocks or flows for which estimates are required.

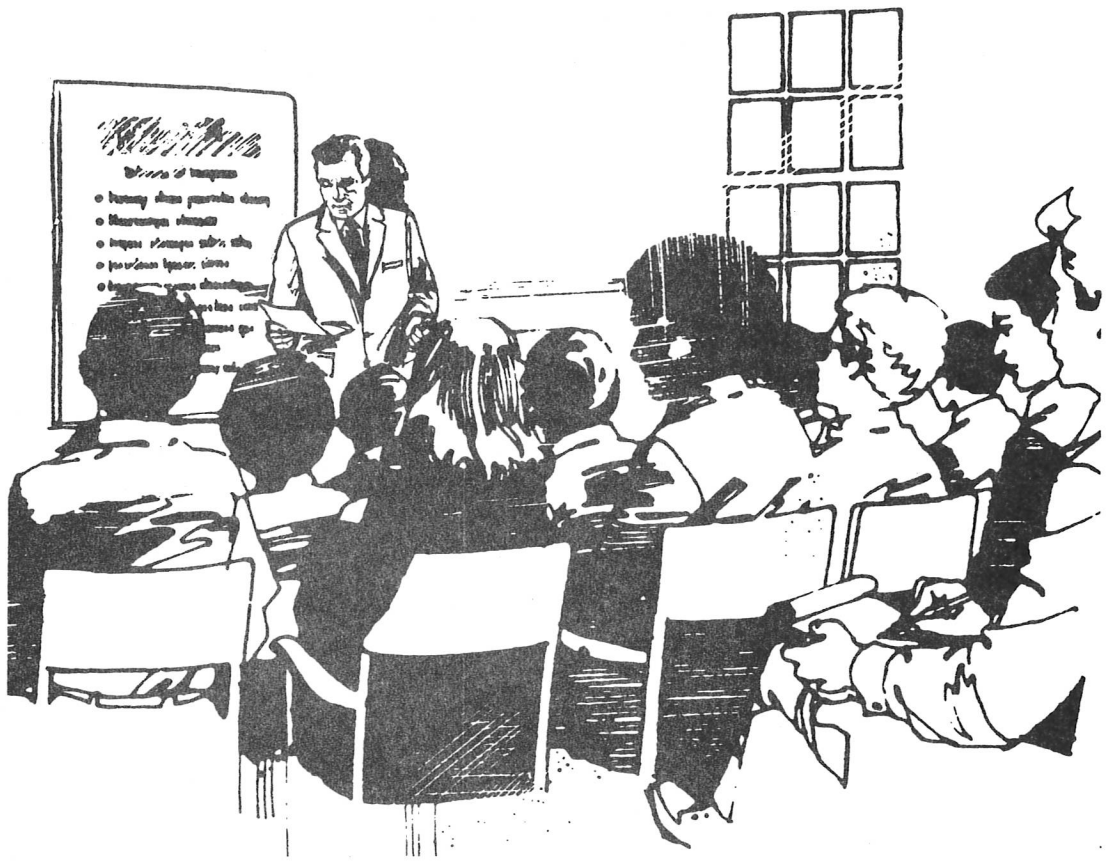
statistical development
1 'back of the envelope guesstimates'
2 analytical estimates
3 direct surveys

This difference in character is one of degree only. The tendency is for the analytical work in national accounts and related economic-model building to inspire developments in the systems of direct collection. So, directly collected data are constantly replacing the often initially rough but progressively refined analytical estimates.

An illustration of this, from Australian experience, is the farm cost estimates (required for the estimation of gross farm product and farm income). These have been produced analytically using a variety of data sources and assumptions (e.g. applying, to acreage under crop, a survey-based standard cost per acre for a particular item of cost). This work led to the establishment of a periodical comprehensive survey of farm expenditure to enable much of the costs to be estimated directly.

Projections

While historical series enable the impact of policy initiatives to be analysed retrospectively, the ultimate in the analytical production of statistics is the projection into the future. Most uses of statistics are in respect of decisions relating to the future and users extend the available data forward to the period which concerns them. They may do this by crude intuitive methods but they do it nevertheless and make their decisions accordingly.



speculation?
or statistics?

Many government agencies are involved in frankly speculative projections of their statistics. However professional the econometric methods employed, official statisticians are inclined to enter this field only on an *all care, no responsibility* basis and would not glorify the forecasts as *statistics* — certainly not *official statistics*. **Nevertheless, projections are in essence no different from those official analytically-produced current estimates which are based on projection of past data and known relationships.** Certainly, involvement in this work has considerable payoff to statisticians seeking to improve their methods of producing current estimates and they can protect their reputation for objectivity by publishing the raw data, assumptions and methodology along with the projections.

Evaluation of accuracy in analytical processes of producing statistics

A characteristic of analytically produced data is that they are generally information which is engineered to satisfy the direct requirements of users (by adapting available basic data for the purpose).

This has some implications for analysis directed at the measurement of error in order to improve the design of statistical systems. **Since the total statistical system must embrace analytical extensions of the basic**

direct collections, it is wrong to concentrate attention exclusively on the elimination of error in the direct collections.

accuracy versus relevance

Sometimes this concentration on readily measurable survey error can lead to design changes which improve the accuracy of the collection by changing its data item definitions to refer to items which can be more readily reported but are not quite the information required. This may reduce response errors or improve response. But the result may be a reduction in the quality of the data in terms of their application to the user's problem — because of the greater dependence on analytical methods to transform the more positive but less relevant series to the one which the decision-makers need.

Thus there is often conflict between the statisticians responsible for censuses and surveys and the producers of synthetic or derived series because of a failure of the former to consider the total system requirements for both the derived and the primary data.

For example, while it may be necessary to know the business component of farm expenditure on motor vehicle operation, it may only be possible in a particular collection to limit reporting errors to publishable levels on total expenditure (ie combined business and personal costs). The choice is then between trying to get the farmer to make a rough split of his total motor vehicle costs or to make the split by analytical methods that may be even rougher.

In theory, it is possible to attribute levels of errors to each stage of the information collection and decision making process and to use this information in directing efforts to minimize the total errors. In practice, only selected information on accuracy and relevance can be obtained and it becomes necessary to use judgment to make appropriate assumptions to fill the gaps.

to improve? or extend the system?

It is important to weigh the value of each collected series with due regard to its size, variability and significance for economic analysis and to avoid the tendency to improve what is already being done without always counting the opportunity costs in terms of other more necessary jobs not done at all.

Analysis/interpretation to amplify publications

drawing out statistical significance

One aspect of this has already been mentioned. It has always been the responsibility of statisticians to try to guide the users concerning the accuracy of the published statistics. Thus, minimal assistance to the users will be a description of the concepts and methods of the inquiry and an evaluation of the accuracy of the estimates.

highlighting

The second aspect is the provision of supplementary analysis to extend the information content of the published figures. This may take the form of a commentary which distils the highlights of the statistical results and expresses them in a literary form which

makes them more readily digestible by the public. This type of analysis aims to be quite objective and policy neutral and yet be as helpful and interesting as publication deadlines will allow.

explaining movements

A more venturesome analysis will go further to explain changes in movements and relationships in the light of associated events. This requires familiarity with the data and a more complete grasp of relevant theory and the general institutional and other factors at play. The process of writing-up such analysis provides a highly desirable discipline and stimulus for exploring more deeply the validity and significance of the statistical results.

occasional papers may offer research findings rather than 'official statistics'

A similar but more general type of interpretative analysis involves a sifting and synthesis of large masses of data derived from censuses and surveys. Monographs undertaken by the statistical agency or by expert users have important benefits in that they help to clarify difficult conceptual, methodological and presentational issues and provide a vital feedback for planning future statistical programmes.

Analysis/interpretation by users

So far we have viewed analysis/interpretation from the perspective of the producers of statistics in a specialised statistical agency.

In a sense there are producers and users at every stage of the statistical production process within a national statistical system as we have viewed it.

But the perspective of users outside the specialised statistical agency may be very different. They will generally not have a reputation as purveyors of 'facts' to inhibit them from drawing inferences.

In fact, if they are to use statistics to solve problems, they are obliged to draw inferences as best they can from the available data. Essentially they will use the same kinds of methods, carrying on from where the official statisticians leave off, to provide answers to their specific questions.

Access to statistical information and services

We have already noted that the distribution of the NSS's professional analysts and interpreters may differ between countries, particularly to the extent that the responsibility for the different subject-matter fields is centralised or decentralised.

So the users of statistics need to know something about the institutional arrangements in different countries for producing, publishing and analysing

statistics, as well as something about the different service philosophies and facilities that they will encounter.

In Australia, users have only to contact the office of the Australian Statistician in Canberra or any of the State offices in order to have an entree to the full range of official statistics produced in that country. This is because Australia has a highly integrated and centralised system of official statistics concentrated almost entirely in the one agency.

Some Commonwealth countries such as Canada and New Zealand have centralised systems, as in Australia.

Others, like India, are modelled more on the British system. Britain has a Central Statistical Office which manages the *statistical service,* has some coordinating functions and produces business statistics and the national accounts. But the system is substantially decentralised — the Department of Trade and Department of Employment are big producers of official economic statistics, the Office of Population Censuses and Surveys conducts the population census and so on.

In the United States there is even more decentralisation. It has a big concentration of effort in the Department of Commerce with its Bureau of the Census and its Bureau of Economic Analysis (which produces the national accounts).

But there are many other central departments in the US, such as the Department of Agriculture, the Social Security Administration and the Internal Revenue Service, publishing large volumes of statistics. At the pinnacle is a small office of statistical policy which seeks to establish some coherence in all this but does not itself collect statistics. Its task is difficult, because of the number of different agencies producing and using statistics and because the system is repeated at the level of the States, (statistical co-ordination is a perennial problem for conferences of State governors).

Of course, there are echoes of the US system elsewhere — postwar Japan, for example, rebuilt its statistical organisation on the United States model.

Two problems arise in obtaining statistics from countries with decentralised systems — (a) it may be difficult to find out just what is available (b) there is a likelihood that the statistics available from different agencies may not be comparable.

Particularly in cases where users wish to make international comparisons it is often advisable to go direct to the publications of the United Nations Statistical Office

and specialised UN agencies such as OECD, FAO, ILO, IMF and UNESCO. In their publications, effort is made to present statistics on a common basis or at least to indicate when a country's figures differ from the standard.

But of course, there will be occasions when users will need to go to the source for the most up-to-date and detailed statistics of the country in which they are interested.

As a general guide, serious users might —

suggestions for the serious user

(a) identify the major statistical agencies;

(b) obtain an index of publications and related guides to services from the central statistical agency;

(c) get on the mailing list for publications which indicate current developments in the national statistical service (eg *Statistical News* in the United Kingdom and *Infostat* in Canada);

(d) get to know any points of ready reference (such as the central information service and library of a central statistical agency, the national library, or the government publications office) and any public access facilities, such as videotext services, Canada's CANSIM database and Australia's AUSSTATS data service;

(e) get on the mailing lists for publications required regularly. Some statistical series can (by previous arrangement) be delivered automatically and instantaneously to the electronic mailboxes of subscribers. Watch for new releases and monitor the media for published news items and detailed articles and reports;

(f) make use of the telephone to ask directly for the information they need (or think they need);

(g) establish working relationships with experts in the NSS who specialise in their field of interest; and

(h) become familiar with commercial intermediaries offering extended data bases and analytical services covering both general and statistical information. The AUSSTATS on-line data service, for example, incorporates the facility to transfer data to personal computers. Access to data by direct electronic means will certainly greatly expand the amount of detailed information which can be made readily available and in a form in which it can be directly worked on.

The selection process

Users of statistics should not be diffident about indicating how they intend to use the statistics. The statistician may then point out the possible pitfalls. If what users want is not available, the Statistician might offer slightly different information that would be just as useful — or something very different to serve as a proxy with some restructuring of the user's analytical model.

Statistical yearbooks often serve as a framework that is helpful in finding the way to appropriate statistics. If the users' interest is in economic statistics, look first at such structural statistics as are found in published national accounts.

Publication patterns

Users will need to find out the scope and general pattern of publication arrangements. For example, the sequence in Australia is:

(1) *duplicated press statements and bulletins* — quick release, short historical series, often issued first as a preliminary estimate, with subsequent revised bulletins;

(2) *printed summaries covering broad fields* of economic indicators — eg an Economic Indicators Bulletin

(3) *printed reports on a particular field with full detail in an historical series* — e.g. a Demography Bulletin;

(4) *comprehensive reference works* — e.g. a Year Book, (textual information as well as tables) or an annual national accounts publication.

Users will also need to know the *revision cycle* in which preliminary estimates are subsequently revised with later information. You may have to make up a long-term series by reference to a sequence of publications, always preferring the figures from the latest publication as far back as they go. In the case of the national accounts, users will need to be prepared for revisions going back five years or more.

Feedback

a crucial need for a healthy NSS

Any system needs some kind of monitoring of its output, feeding back to control the system. This is particularly so for a national statistical system. Continuous thoughtful communication between users and producers of statistics is essential to signal the need to adapt the system to changing national requirements for statistical information.

Formal channels of consultation are needed to ensure that priorities for the development of programmes and projects are allocated in an optimum way. These need to cover potential as well as existing users, in order to identify and meet new information requirements and to counter the natural tendency to concentrate on improving existing series.

Users need to invest time and effort in this regular or periodical process of review and redevelopment to ensure that the information system they get is the system they need.

To sum up

We are back where we started on our guided tour of the processes involved in a national statistical system. We have seen the possibilities of computer based processing, archiving and analytical techniques. But we have been reminded that the organisation of mass social data, as an instrument of national-level research, policy, planning and management, involves much more than the loading of large computerised data bases for users to access and manipulate.

The task is to develop and continuously adapt information systems. The systems are interdependent and complex. Evaluation of their effectiveness is an essential step in the continuous process of data gathering and management for purposes of analysis and interpretation in the context of continuing information system development and redevelopment. Necessarily, responsibility for the system has to be shared between the providers of basic data, the user of statistics and the government statisticians as intermediaries.

PART 3: PRODUCTS FOR THE ECONOMIC ANALYST

CHAPTER 10
What's so distinctive about economic statistics?

Statistical products and subject-matter divisions
So, what **is** so distinctive about economic statistics?
National accounts as a coordinating framework
What are the boundaries of the accounting system?

CHAPTER 11
Economic accounting design — bridging accounting and economics

General objectives of national accounting
Balanced national aggregates — GDP, national income and all that!
The accounting format and economic identities
Relating flows to stocks

CHAPTER 12
Economic accounting design elaborations

A more detailed framework
Real and financial perspectives
Two ways of dividing the economy into sectors
Input-output analysis
Flow of funds
Domestic territory and resident economic agents
Rest of the World and Balance of Payments
Adaptations to country circumstances and needs

CHAPTER 13
Extensions beyond the central framework

New pressures from changes in social priorities
What are the system dimensions of the accounting approach?
Extensions and elaborations of national accounts for broader social purposes

CHAPTER 14
The economic accounting process

Alternative accounting routes
Economic transaction flows — points of observation
Confrontation of data
From the bottom up? or from the top down?
Input-output tables in the accounting process
GDP at constant prices
Balancing items, residual discrepancies, accuracy and reliability
Beyond the national income and expenditure accounts
Concluding reflections on economic accounts and processes

CHAPTER 15

Primary and special purpose derived economic data systems

Broad categories of economic statistics
General problems of system integrity
Characteristics of access routes to source records
SNA's presumptions about statistical units and their records
Problems for a national economic statistics report system
Real life units and standard statistical units

CHAPTER 16

Systems based on enterprise-type owning units and institutional sectors

Institutional units and sectors
Nonfinancial corporations sector — enterprise statistics
Financial corporations collections
General government collections
Private nonprofit institution collections
Household sector
Rest of the world

CHAPTER 17

Elements of systems based on establishment-type producing units

Basic production process concepts
Registration and classification of establishments
Nature, purposes and principles of industry classification
The assessment and application of an industry classification
The geographical dimension and regional accounts
Statistical programme groupings
International trade statistics and commodity data

CHAPTER 18

Main fields of establishment level statistics

Agricultural statistics
Industrial statistics — mining, manufacturing, electricity, gas and water supply
Statistics of distributive trades and related services
Statistics of construction
Statistics of transport and other services
Government services
A last word on economic statistics

CHAPTER 10

WHAT'S SO DISTINCTIVE ABOUT ECONOMIC STATISTICS?

In Part 2 we set out to identify the elements of a system in the nation's arrangements for producing and using statistics. In Parts 3 and 4 we shift our examination from the process to the product. We try to discern what is systematic about the composition of the statistical data which are produced. We will try to identify the major data systems and subsystems likely to constitute an NSS, concentrating on what makes them distinctive and on how they can be linked through elements they have in common. In Part 3 we focus on so called 'economic statistics'. In Part 4 we examine 'social statistics', with particular reference to noneconomic data.

Statistical products and subject matter divisions

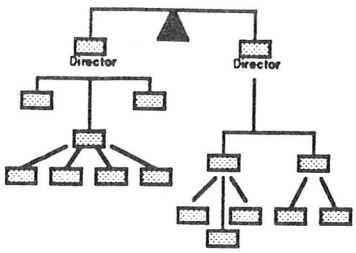
a systematic process — and product

We are familiar enough with the notion of processes being formed into a system — as when we link a radio, tape player, amplifier and speakers to form a sound system. **We are perhaps not so familiar with the idea that the product of statistical information produced by a statistical service might constitute a system** by virtue of the consideration given to ensuring that the various separate statistical series produced will hang together in systematic ways.

We have seen how general purpose national statistical systems can be visualized as the tip of the iceberg, kept afloat by the mass of data systems which provide accounting and other operating information within the individual enterprises and institutions of society, primarily for use by the organisations themselves.

In this sense, a *national statistical system* relating to the social and economic affairs of a nation is an all-pervasive one and could cover almost any activity and resources of human society. We will limit our *guided tour* to the major socioeconomic fields which central statistical agencies are commonly involved in coordinating. In the context of producing (say) an *Official Yearbook* they will spread their interest to touch on political attitudes and physical science data, such as voting behaviour or meteorological records, but only to the extent that they help readers to understand the nature and state of the society.

Within this near universal range of subject matter, central statistical offices generally draw a distinction between *economic statistics* and the rest of statistics about society (which is likely to be labelled *social and demographic statistics* or *social, demographic and labour statistics*).

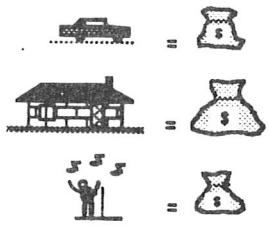


Each of these broadest categories may in practice be represented by a number of *subject matter* divisions of the statistical organisation. Such groupings of responsibility may reflect an effort to define a hierarchy of levels of data systems corresponding to recognisable real world systems. But internal considerations, such as the need to even up the work loads of divisional heads, may make this difficult to achieve.

So, what is so distinctive about economic statistics?

One can create more or less integrated combinations of statistical fields, such as separate birth, death and marriage statistics being designed as subsystems of a system of vital statistics which, in turn, is part of a system of social and demographic statistics. This might be pursued by making the most of standardising such common elements as the statistical unit (eg the household or family), source documents, or processing procedures, using a common publication vehicle or by making the programme the responsibility of a particular agency or unit.

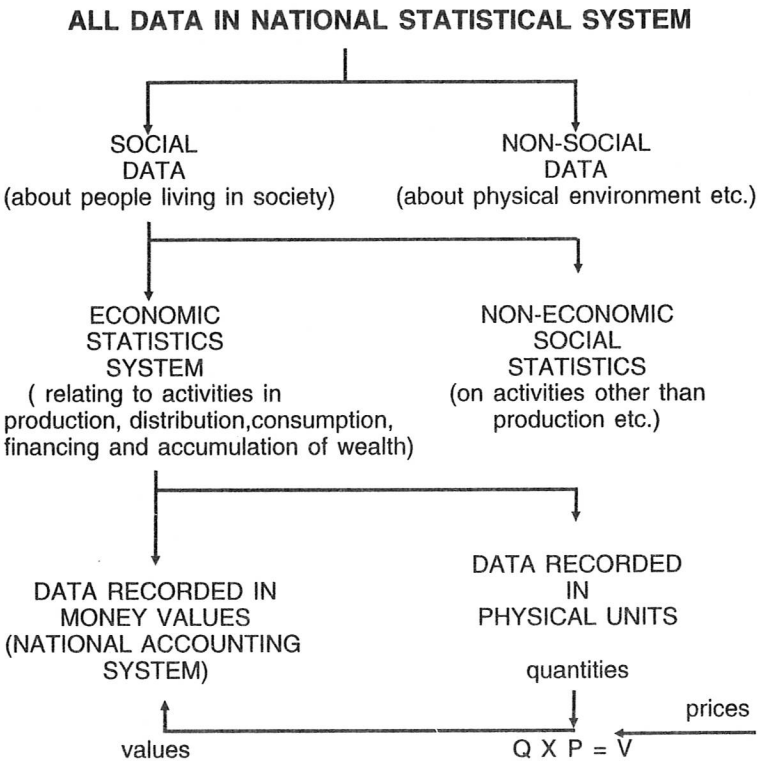
The great range of primary national economic data systems is remarkable in the extent to which they have been associated to form broader systems by this kind of process. This is so because all the activities, transactions, products, stocks and flows covered in these data systems are capable of being accounted for in terms of the same numerical values.



Money values determined between buyers and sellers on the market are the common measuring rod. All transactions of any kind valued in this way can thus be added up to provide national aggregates.

This means that the basic economic processes of producing goods and services, distributing income, consuming the product, financing these activities and accumulating wealth can then be depicted as a connected system, with regular interaction and interdependence, because all these activities in a given country can be measured in the same dollar terms.

We can picture the economic statistics system in relation to the total national statistical system as follows:



While it is possible to identify data systems readily enough as being economic data systems because they relate to the basic economic processes, those relating to production and consumption can be (and may need to be) analysed in terms of physical units as well. As the diagram indicates, transformations between physical units and the money value units of an economic accounting system can be made, given appropriate prices, since quantity X price = value.

The economic analyst will generally work with the accounts and supporting tables of economic accounting and primary data collection systems which record transactions and summarise results in money value terms.

Analysts who are not primarily interested in the economic process may also use these data. They may need (say) data on public order and safety expenditure in assessing the security of the community, even though the concept of security itself is not one that can be evaluated in market terms. By the same token, the compilers of national income and expenditure estimates may need to draw on population census demographic data as an input to their calculations. In short, while data are not intrinsically economic or noneconomic, data in money value terms are the natural material for the analysis of economic processes.

National accounts as a coordinating framework

As the previous diagram indicates, national accounts provide the general framework for statistical systems which focus on economic processes and on money values as the unit of account.

SNA — key to integration of economic statistics

By reason of their comprehensiveness, standards set in the United Nations, *A System of National Accounts* (SNA) become the basis of (or at least the point of departure for) the standards of particular fields of economic statistics.

For historical and other reasons the total body of systems of economic statistics may be far from being a fully integrated system in any country, but the basis for this is well developed in the SNA and related manuals, including expanded and updated manuals on sources and methods and reviews of various kinds pursuing the development of the system.

Why is SNA so influential?

There are two basic reasons for the strength of the SNA in its coordinating role. **Firstly, it applies the concepts and conventions of business accounting** to the task of summarising transactions and outcomes at every level up to that of the nation as a whole. **Secondly, its design reflects theoretical concerns** to provide explanations of economic behaviour and the basis for reporting, predicting and influencing this.

The basis in accounting conventions (with necessary allowance for the special perspectives of economic agents which are not engaged in business activity) **anchors the economic statistics system in the solid ground of basic transaction record keeping and primary data collection systems.**

The recognition of economic theory in the design of national accounts helps to ensure that the statistics are relevant enough to induce a lively interaction with those developing theoretical explanations and models and formulating economic policy.

Taken together, the individual transactions and economic theory attributes of the SNA ensure that the SNA is a **general purpose, open ended, and flexible framework for economic data systems and analyses.**

MICRO/ MESO/ **MACRO**

In fact there is no necessary distinction to be drawn between national accounts and the rest of economic statistics. They essentially represent different ends of a continuum of aggregation in which one moves from the accounting perspective of individual self-seeking enterprises, households and the like (*micro data*) to the economic accounting perspective of larger groups, such as industries, institutional sectors and regions, up to that of the nation as a whole (*macro data*). (The term *meso data* has some currency in the literature to represent an intermediate level in this continuum.)

SNA's changing perspectives

As an expression of economic theory, the SNA has developed from the initial concentration of economists and national accountants on the **outcome** of the economic process (in terms of concepts such as gross

domestic product and national income which are seen as objects of maximisation for a country). This was elaborated in the initial 1953 SNA as a statistical description of the **process as such** under the strong influence of the traditions of Keynes' (1936) *macro economics.* This gave prominence to a number of key abstract variables relating consumption, savings and investment to income, expressed in money values and relating to the economy as a whole.

The 1968 SNA took into account additional intellectual streams represented by the input-output studies of the production process and the variety of process oriented studies of financial flow, capital financing and income distribution, which require detailed breakdowns for different categories of transactors in the economy and for different kinds of transactions.

While claiming to be neutral as to economic models, the SNA tries to accommodate a range of important views of the economic process. The SNA has both supported and stimulated a great range of models for analysing national and regional economic growth and change.

The SNA Revision 4 (probably to be published in 1994) will elaborate and extend the scope of national accounting analysis in response to new pressures from changing social priorities such as concern for the environment and conservation of natural resources.

What are the boundaries of the accounting system?

The previous diagram indicated that economic statistics relate to the processes of production, distribution, consumption, financing and accumulation and that an economic accounting system records this in money value terms. The money values are established in the market as we buy and sell goods and services in the course of our economic activity.

scope of accounting system determined by what we include as 'production'

The accounts which are kept by economic agents, or at least are in their minds (in the case of households), will relate to —

- the net value of what each enterprise *produces*;
- the *income which arises* from this;
- the *consumption, saving and investment* out of the distribution of this income;
- the *financing* (lending or borrowing) of net deficit or surplus positions and

- the *change in the net worth* of the producer.

Such summaries of transactions and outcomes, analysed in various ways, are fundamental to decision making and planning at all levels.

Thus production generates the incomes from which the investment for economic growth and the final consumption of goods and services of the people can be financed. What is then covered by economic accounts and supporting data systems will depend on what we define as *production* and on the transactions we are able to record on this basis.

The *boundaries of production* to be included in national accounts have been established in the SNA as a compromise between the practicability of coverage and the need to explain economic behaviour comparatively over time and between countries.

The boundaries have not been rigidly set for all countries and there is a natural pressure from the users of statistics to extend the boundaries of the *core* SNA system or to widen the scope of *production* by way of supplementary analyses.

drawing the boundaries — the marginal areas

The fringe areas are those in which the production of goods and services takes place without market transactions or where the transactions are of a clandestine nature and so are difficult to cover directly. In brief, the SNA Rev. 3's guidelines are as follows:

Illegal activities — include only activities in which the receipts are obtained with the unenforced consent of the payer — eg smuggling and distribution activities in the black market but not the proceeds of extortion. This, of course, is a statement of principle. What happens in practice is less positive — the drug pedlars' production of services may not be reported directly but their income and consumption arising from this may well enter into the accounts indirectly, perhaps with a discrepancy between the income and expenditure recorded for them.

Production for own use — the 1968 SNA lists activities which are to be included whether provided for the producers' own consumption only, or produced both for their own consumption **and** for sale. These are —

(i) primary production,

(ii) processing of primary products,

(iii) production of fixed assets,

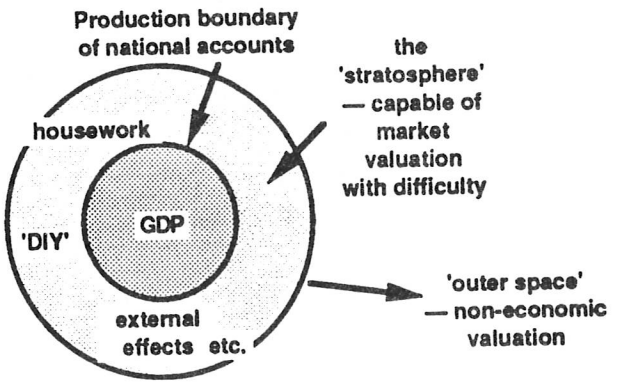
(iv) provision of housing services (enjoyed by the owner-occupier of a dwelling).

(v) production of **any** other goods and services where the producers also sell some of their own output.

The SNA Revision 4 will follow a different approach to defining non-marketed production to be included, but effectively the changes will be minor.

Instead of the listing above, the SNA Rev. 4 draft offers a general rationale for defining the boundary of production for its fully articulated central framework — *'This production covers all goods and services which are actually delivered to other economic units whether paid for or not; it covers also goods which are retained by the producer for his own use but which could have been offered to other economic units on the market; finally it also includes services produced for own use by processes of production in which the factors of production employed are remunerated by the producer. Thus services rendered by government to other units free of charge or at a nominal price are included in production, as well as services destined to collective consumption of society. On the contrary, domestic or personal services rendered by members of a household for consumption within the same household are not included within the production boundary of the integrated system'.*

Thus, significant areas of production for own use which are not covered in the SNA's central framework (both in Rev. 3 and 4) are housework and other *do it yourself* (DIY) activities that are undertaken in and around the house by householders. These are certainly productive, although difficult to evaluate in credible money value terms. Beyond this are many other activities productive of welfare, such as voluntary work with non-profit organisations.



Generally the view has been that conventional boundaries have to be drawn in some standard way if firm comparisons of production and welfare aggregates and growth over time and between countries are to be made at reasonable cost. Extending the production boundary beyond recorded market transactions necessarily involves difficult data collection and highly conjectural valuation.

This is particularly the case for activities for which the product is services, rather than goods. This is why it has been considered advisable to treat studies of the value of household work, *do it yourself* (DIY) activities and the like as being supplementary special-purpose extensions of the national accounts and to omit them from the standard national aggregates, such as gross domestic product.

International comparisons of different estimates suggest that, in the case of even a highly developed market economy, an imputation for unpaid household work, for example, may amount to anything from 30 to 60 percent of GDP as currently defined, depending on cultural differences between countries and on which necessarily arbitrary assumptions are adopted. Economic analysts argue that if this kind of activity is incorporated in the system's central framework it will detract from the accounts' primary function of displaying the interactions of different sectors of the economy with one another in terms of the market transactions in which they engage — obscuring basic economic analysis in terms of prices, output, and employment.

Household Work (Allen and Unwin Ltd, Sydney, 1989 edited by Duncan Ironmonger,) summarises the state of much of this research, and of the related conceptual and policy issues.

Generally, the need for consistency has been the paramount concern and a limited but reliable coverage of economic activity has been accepted as serviceable enough if it can be assumed to reasonably represent the broad picture of secular progress in the basics of quality of life.

The general problem is that the more ambitious the resort to imputations the less the credibility of the accounts as an instrument of economic planning and management and a basis for international comparison. Certainly the current review of the SNA indicates that many users would prefer to focus sharply on the behaviour of the market economy, at least in the first instance, since it is this activity that is of primary concern for macro-economic policy and on which most policy instruments bear.

Accordingly, the review SNA considered that 'household services (should) continue to be excluded from the production boundary because they are not capable of being marketed once produced and the quantity produced is not related to market forces as directly as in the case of goods' However it has recognised that to address many areas of interest 'it will be appropriate for separate analysis to be undertaken with production boundaries defined much more widely. It is expected that such analyses would also be undertaken in the form of specialised satellite accounts'. International Association for Research in Income and Wealth, 21st General Conference, Lahnstein, 20-26 August 1989, 'Major Changes Proposed for the SNA: An Overview'.

The need for and possibilities of exploring activities beyond the standard production boundary may well change and will vary from country to country. The ABS, for example, supports the development of such accounts as part of the revised SNA, and proposes to work towards the development of satellite accounts for Australia incorporating estimates of the value of unpaid household production. Their timing, frequency and quality will depend on the availability of appropriate source data and the resolution of what are accepted internationally as very difficult methodological issues.

So, what do we conclude? The fact that the *core* UN national accounting system cannot embrace all aspects of human condition in a single indicator does not negate its value as an instrument for monitoring a very significant part of our activity, provided that we are fully aware of the positive and negative elements of social progress that it does not account for.

But, in so far as there is an expectation that standard aggregate net outcomes, such as GDP, will be a sufficient index of the progress of the economy, the system has serious shortcomings. In Chapter 13 we will consider some of the efforts to extend the core system to take account of aspects which can be explored in market value terms for more thorough social analysis.

To sum up

We set out to answer the question 'what's so special about economic statistics?' 'What distinguishes them from other social statistics?' The key, we decided, is the common unit of value. The economic accounting systems remain essentially a record of transactions which actually occur on the market, or which involve decisions to produce, consume etc which are in the nature of market decisions and to which market values can be sensibly attributed.

The strength of the system is its universality and objectivity. The values are those which are applied in the actual decisions which govern the economic processes.

We may not make all our decisions within this market framework (and sensibly so, since market valuation of production cannot adequately account for (eg) negative environmental effects of production, or objectionable moral implications and other extramarket considerations). But such an accounting system does at least marshal a great mass of information consistently within the framework of economic decision making that is generally familiar to individuals, enterprises and governments. Subsequent chapters will explain how this has been achieved.

CHAPTER 11

ECONOMIC ACCOUNTING DESIGN — BRIDGING ACCOUNTING AND ECONOMICS

We have considered the nature of economic statistics and the boundaries of economic accounting in relation to the total body of statistics about society. We now focus on the United Nations' System of National Accounts (SNA) as a set of guidelines for designing national accounting systems that will enable the activities taking place within the conventional production boundary to be summarised for a great variety of analytical purposes.

In this and the next 2 chapters we concentrate on the WHAT and WHY of economic accounts before looking at the HOW of the accounting processes in Chapter 14. At this point it is enough to note that a system of economic accounts for any particular country evolves as much from sources and methods development as it does from the aspirations of the designers and users of the accounting system. But a great deal of momentum certainly arises from the existence and general acceptance of the SNA as a world design standard, adaptable to the circumstances and needs of particular countries in different stages of development.

Centrally planned economies, such as the USSR, have used a different system in the past, called the Material Product System (MPS), but there are many common elements and some countries, such as Yugoslavia, have prepared estimates on both bases. Most of the centrally planned economies are now moving to adopt the SNA framework in conjunction with their adoption of market or mixed economies.

As we have already noted, it is fundamental that the developing economic accounting systems should be designed to assemble information in ways that will support the main thrusts of theoretical explanations, forecasting and policy models of economic processes. We will look at the elements of the SNA design with this in mind.

General objectives of national accounting

At the time of writing this handbook, the SNA was close to finalisation as Revision 4. A draft overview of the System puts the general objectives very clearly —

> *"A system of national accounts tries to give a complete description of the economic system. It analyses the main economic functions and their inter-relations. It shows how production creates or transforms goods and services and at the same time generates income, which is subsequently distributed and redistributed; then it shows how income is used for present final consumption and for accumulation which will permit raising consumption in future periods; it also depicts the financing of the economy through the use of a number of financial instruments and the role of financial intermediaries."*

The paper goes on to add that in real economic life these phenomena are taking place simultaneously; economic agents try to anticipate the different phases of the economic process and more generally the linking of various periods of time.

Continuity of the economic process over time means that flows which occur in a given period and stocks that exist at a certain point of time are closely related. Thus in their description of the economic process, national accounts

aim at covering the analysis of wealth, through balance-sheets, as well at that of flows. Moreover, as the economic process involves relations between the economies of different countries, external transactions are naturally dealt with by national accounts.

Balanced national aggregates — GDP, national income and all that!

GDP — an object of maximisation — pivotal in analysis of activity

The interest in calculating a single comprehensive *economic welfare* aggregate has a long history and countries still set national goals explicitly or implicitly in terms of levels of Gross Domestic Product (GDP) *per capita* or rates of growth in GDP (generally after the effects of inflation have been removed from the dollar values).

While the UN SNA is not simply a system for generating internationally consistent estimates of GDP, **the concept of GDP is pivotal to the SNA's representation of economic activity**.

Production is the outcome (in the form of goods and services) of the process whereby labour, accumulated capital assets and technical knowledge are applied to the task of transforming natural resources and semi-processed goods.

product— a geographical concept

The standard UN measure of this, the *Gross Domestic Product,* **summarises production of a geographical area — ie within the country's domestic territory.** The system of accounts then shows (eg) how different industries have contributed to this product, how incomes arise from payments for labour and other factors of production and how the product is consumed, exported or is added to capital equipment or stocks for use in production in the next accounting period.

In fact the GDP is not simply the *pile* of goods and services produced during the period, because much of this output is used as input during the period.

Gross Domestic Product is net output, but is gross of depreciation

Thus, although it is called *Gross Domestic Product*, it is a net output concept in that it is the total market value of goods and services produced in a country, after deducting the cost of goods and services used up in the process of production. It is *gross* only in so far as it has not netted off depreciation.

In principle, to obtain the true *production* (ie net domestic product), one should also deduct the value of capital depreciated or *used up* during the period. However, because there is no way to directly record capital consumption and no unique way of imputing it, the concept of gross domestic product has been preferred by the UN for its key international comparisons as being more reliable than net domestic product.

In Chapter 12 we will be looking at this concept of GDP in terms of accounting origins and we will see how it can be derived from data on:

- the gross values of production and operating costs of producers; or
- the incomes received by owners of *factors of production** ; or
- The expenditure on GDP, ie on the *final* goods and services (net of imports), which are bought during the year by residents or exported, excluding those used up in the production process as *intermediate consumption*.

* *Factors of production are land, labour, capital and entrepreneurship and the corresponding factor incomes are rents, wages, interest and profits. Note that they do not include transfers, such as gifts.*

income produced and income received

These alternative approaches simply reflect accounting equivalences expressed in the system of definitions. Thus, in the simple model of the world as a whole (or for a Robinson Crusoe type closed economy) production, income and expenditure are equivalent by definition. However, for individual nations involved in international trade and financing, they will not be equivalent.

The SNA specifies a series of key related aggregates which bring out the fact that the product of a country's domestic territory may accrue in part to nonresidents who have invested capital in the country or have supplied labour or other services to it. Income *produced* in a country is thus distinguished from income *received* by residents of the country.

3 faces of GDP

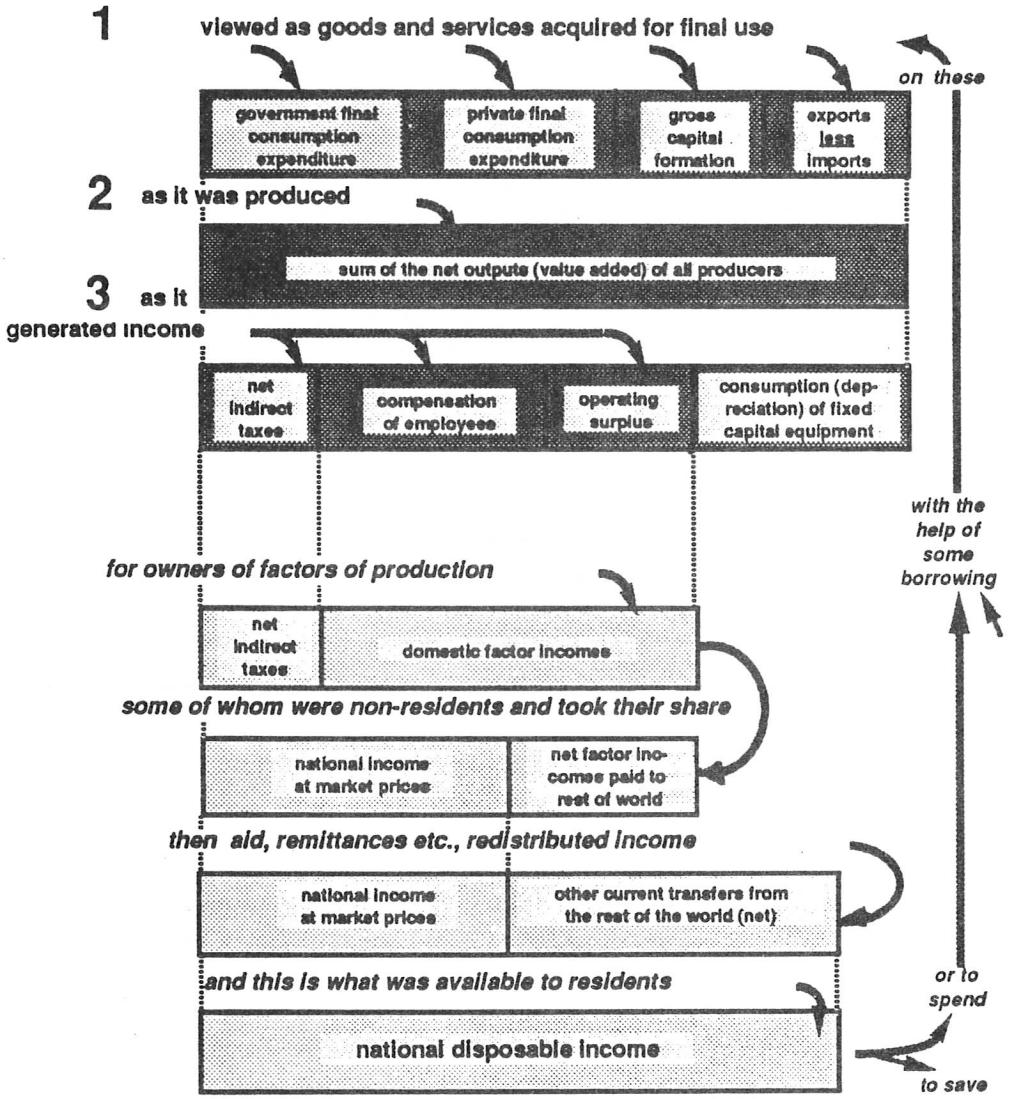
In the next diagram, GDP is shown in terms of —

firstly, the domestically produced and imported goods and services acquired for *final use* by residents and the rest of the world;

secondly, in terms of the aggregate of *net output* (value added) of all producers in the country; and

thirdly, as the *costs* incurred by producers, part of which consists of the incomes arising by way of compensation of employees and operating surplus (which may be distributed to owners and investors).

3 FACES OF GDP



These domestic factor incomes are shown in this example of a hypothetical developing country as being largely payable abroad, (ie domestic factor incomes accruing to overseas owners exceeds factor income accruing to residents from their ownership of factors of production employed overseas). As a result, in this example, national income is much less than the size of the GDP might lead one to expect. But in this case grants etc more than restore the loss to overseas of 'income produced'.

The amount available to spend or save after this sequence of international transfers is entitled *National Disposable Income.* In principle, this would be a better measure of economic welfare than GDP. However, until depreciation estimates have achieved respectability, international comparison is likely to continue to focus mainly on GDP.

Some countries and agencies (notably the World Bank, USA and Canada) continue to favour the concept of Gross ***National*** *Product which is the income of all of a country's factors of production, regardless of where these factors (eg capital invested overseas) are located. This concept may be useful for purposes of making national (ie resident) income comparisons, without having to deduct the often very doubtful depreciation element. But it is not an easy concept to grasp, being essentially an income rather than a product concept. Also it cannot be directly related to production by industry of origin which a country can only monitor and control effectively within its domestic boundaries.*

In SNA Rev.4, GNP is to be more aptly renamed Gross National ***Income*** *(GNI).*

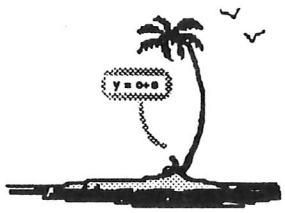
The accounting format and economic identities

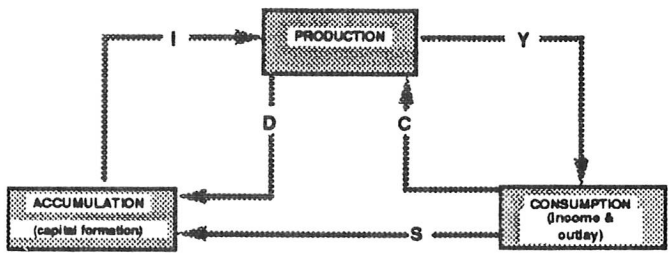
Associated with the identification of such key national aggregates as discussed above has been the development of highly aggregated models of the economy, which represent the dynamic relationship of the main aggregates associated with economic activity — output, consumption, savings and investment and so on. Correspondingly the SNA's accounting model is designed to quantify and systematically relate these key stocks and flows.

The basic format of the SNA's accounts derives from the concepts laid out by Keynes in his *General Theory of Employment, Interest and Money* (1936). They thus conform with the post Keynesian view of a Western style, modern *mixed* economy in which profit oriented private enterprise is guided and restrained by a government which also carries out a considerable amount of economic activity.

The accounting format, which shows payments on one side and receipts on the other, lends itself naturally to expressing the macroeconomic identities. Thus every economic flow is represented as a payment from one account and a receipt into another.

To illustrate, if Robinson Crusoe were familiar with Keynes' *General Theory* and wanted to analyse his economic activities by keeping separate accounts of the producing, consuming and accumulating activity on his island, he might come up with a systems diagram like this:





This would picture the island's production account as a box in which Robinson has to pay for labour services and (hopefully) is returned a profit on his investment in the means of production.

This production gives rise to an outgoing of **Y** (income) from the production account which is also an incoming to the consumption (income and outlay) account.

This, in turn, will show how much of the income is spent on consumption goods and services, which is a payment **C** from the consumption account and receipt by the production account.

The balance is a flow of savings **S** which, with the amount set aside for consumption of fixed capital (depreciation **D**) provides the funds for purchasing new capital equipment **I**, a flow from the accumulation account into production.

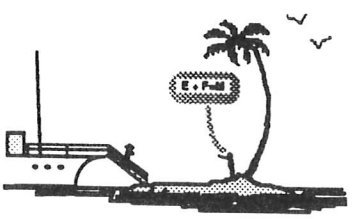
It is important to note that the same economic agents may appear as both producers and owner/consumers. Separate accounts identify their activities in different roles.

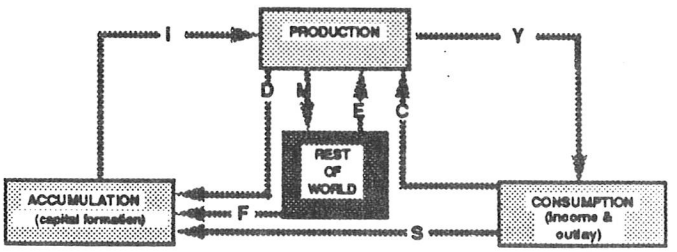
These are **money flows.** Goods and services and factors of production can be pictured as flowing in the opposite direction to the payments and receipts. Thus the payment **C** by owners/consumers to producers is for goods and services for final use flowing to them from producers.

These counterpart physical flows are recorded in *input-output* tables, which involve studies of inter-industry flows and of the flows of particular commodities — ie showing which goods and services are transformed into which other ones. The point to note is that they relate to the same transactions as appear in production accounts, but they focus on the **physical flow** side of the transactions.

However we are concentrating now on the incomings and outgoings of the standard economic accounts, analogous with business accounts, which are presented as monetary transactions.

Suppose the island starts trading with the rest of the world. Adding the flows M for imports and E for exports to the rest of the world, we can elaborate our diagram as follows:





Each of these flows between boxes should be understood as a payment from one account matched by a receipt in another.

Thus, when Y = income, M = imports,
D = depreciation, C = consumption, I = investment, E = export, S = savings, and F = foreign lending, we have —

	Payments (outgoings)	*Receipts (incomings)*
Production account	$Y + M + D$	$C + I + E$
Consumption (income and outlay account	$C + S$	Y
Accumulation (capital finance account)	I	$S + D + F$
Rest of the World account	$E + F$	M

Here the Rest of the World's account closes the system. The flow F represents (from the rest of the world's viewpoint) net foreign lending to support the excess of the domestic economy's import payments over export receipts.

The SNA's illustrative **Four accounts of the nation** (as depicted over the page) presents this a little differently. Income is treated as gross of depreciation and appears in the consumption account as well as in the production account.

Also, in the Rest of the World account, current transfers (gifts etc) and factor income payments are distinguished.

Such summary accounts are built up by way of *consolidating* a number of component accounts.

For example, the Domestic Production Account can be arrived at as a consolidation of Commodity and Activity accounts, while the Consumption account is a consolidation of Consumption Expenditure and Income and Outlay accounts and the Accumulation account is a consolidation of Capital Formation and Capital Finance accounts (consolidation simply means adding up the same sides of the two accounts and cancelling out common terms in the process).

The SNA Rev. 4 draws out the accounting structure much further than this to identify useful new concepts as balancing items and elements of national aggregates of major analytical interest

Note that the accounting formats shown in SNA manuals are intended for expository purposes rather than as a basis for publishing national accounting data (which are generally best presented in tables of time series).

We focus on the SNA Rev. 3's simplified **four accounts of the nation** in order to demonstrate the comprehensive nature of the System's balanced accounting of the flows of production, consumption, accumulation, and international transactions.

THE FOUR ACCOUNTS OF THE NATION
Table 1.1 of SNA Revision 3

1. PRODUCTION
(Domestic Product Account)

Outgoings		*Incomings*	
1. Gross income payments (value added) (9)	255	3. Sales of consumption goods (6)	210
2. Purchases of imports (18)	54	4. Sales of capital goods (12)	47
		5. Sales of exports (16)	52
Total	309	Total	309

2. CONSUMPTION
(Income and Outlay Account)

Outgoings		*Incomings*	
6. Purchases of consumption goods (3)	210	9. Gross income receipts from domestic production (1)	255
7. Saving (15)	27	10. Less Provisions for the consumption of fixed capital (13)	-19
8. Net current transfers abroad (19)	4	11. Net distributed factor incomes from abroad (17)	5
Total	241	Total	241

3. ACCUMULATION
(Capital Transations Account)

Outgoings		*Incomings*	
12. Purchases of capital goods (4)	47	15. Saving (7)	27
13. Less Provisions for the consumption of fixed capital (10)	-19		
14. Net lending abroad (20)	-1		
Total	27	Total	27

4. THE REST OF THE WORLD
(Balance of Payments Account)

Outgoings		*Incomings*	
16. Purchases of exports (5)	52	18. Sales of imports (2)	54
17. Net distributed factor income payments (11)	5	19. Net current transfers (8)	4
		20. Net borrowing (14)	-1
Total	57	Total	57

NOTE that the numbers in brackets in the table above indicate the contra entries of the items numbered without brackets. Thus, the item 3 (sales of consumption goods) in the first account has its contra entry at item 6 in the second account (purchase of consumption goods).

As the accounts illustrate, by tracing the contra entries (in brackets) in the accounts, the processes of production, distribution and accumulation are related in the context of international transactions.

This is not simply a matter of bookkeeping. Dynamic elements like investment are identified and an appropriate framework is provided for forecasting and influencing the overall performance of the economy through fiscal policy, public expenditure and monetary and trade policy.

The accounts as a planning tool — an illustration

Thus the elements of the Domestic Production Account have long been of great interest to governments employing what has been called *gap analysis* for economic management purposes.

In an EX POST (after the event) sense, the Domestic Production Account must balance, with expenditure on GDP being equivalent to the cost structure of the GDP.

'ex-ante' (before the event) forecasting

However, if this account is used as a framework for forecasting, the figures need not balance. In this EX ANTE (before the event) sense, the account can be independently projected in the short term, by means of whatever indicators are available and whatever judgement the forecaster may use. If the projected right hand side exceeds the projected left hand side, it may imply that there is an excess of demand for resources over the expected supply, causing either price increases or increased imports, or both.

ex-post (after the event) outcomes

Whatever the mechanism, the account will ultimately balance in the EX POST sense. However, if the government regards these adjustments as undesirable — for example because they endanger price stability or balance of payments stability, then they will have to adopt policies which eliminate the EX ANTE imbalance by other means (eg by reducing public expenditure, or by discouraging private capital expenditure through raising interest rates or otherwise restricting credit, or by increasing taxation to reduce personal consumption expenditure).

Alternatively, if the projected left hand side exceeds the projected right hand side, this will imply an excess of supply over demand for resources, causing unemployment or decreased imports, or both.

Again, this account will balance EX POST as a result of these adjustments, and again, if the government regards these adjustments as undesirable (eg because it endangers the full employment objective) then it will have to adopt policies which eliminate the EX ANTE imbalance by other means (eg by increasing public expenditure, or by stimulating private capital expenditure through lowering the interest rates or otherwise liberating credit, or by reducing taxation to increase personal consumption expenditure).

SNA variables needed for most economic models

Applied in much more sophisticated ways, **economic accounting systems based on the SNA standard are able to accommodate a considerable range of macroeconomic models of the economy, in so far as they have common information requirements.** Thus there is the same concern for variables like operating surplus, consumption, saving and investment whether one is a Keynesian economist or a monetarist.

Developing countries may sensibly have little faith in prescriptions based on the dynamic chain effects of increases in income, stimulating consumption and, in turn, investment and production. They may focus their interest on the chronically low levels of saving and look to stimulus from foreign aid and borrowing and international trade. **Nevertheless basic national aggregates like GDP are still likely to be key variables in their view of the economy and in their particular application of the macroeconomic approach in analysis and policy.**

Relating flows to stocks

We have been considering only the current accounts of the System as summarised in a set of National Income and Expenditure Accounts which are generally the first priority in setting up a system of national accounts. But it should be noted that, as with commercial accounts, the SNA prescribes a fifth major account, the Balance Sheet, which records 'stocks' rather than 'flows' — the Nation's accumulated stock of wealth at a given point in time.

The Balance Sheet is linked directly to the Capital Finance account. Over a given accounting period the 'closing balance sheet' is equal to the 'opening balance sheet', plus or minus the changes in capital stock during the accounting period. These changes are shown in the Capital Finance account.

In Chapter 12 we consider the natural elaborations of the basic national accounts which arise from subdivision of transactor and transaction categories and the integration of such logically contained, but historically distinct, fields as input-output, flow of funds and balance of payments analysis.

CHAPTER 12

ECONOMIC ACCOUNTING DESIGN ELABORATIONS

As the SNA's accounting structure is elaborated, a great number of more detailed transaction flows and balances are defined and classified and these form part of the standards for economic statistics generally.

Then a further dimension of information is added when the accounts are specified for different transactor groupings.

We discuss here the development of national accounts as a basis for detailed structural studies and take in input-output studies, flow of funds and balance sheets and explore the Rest of the World sector and balance of payments analysis.

A more detailed framework

While macroeconomics remain important and vital in various forms, there has developed a great range of theories and models which are both specialised and economy wide. In response, economic accounts based on the SNA's 1968 version have been elaborated and extended to the point where they provide a detailed general-purpose data organising framework relating macro to meso/micro economic analysis and policy.

elaboration by sub-division of transactions and transactors

There is thus a much enlarged range of variables now defined and related in the accounts. At the same time, with the subdivision and disaggregation of variables, it has become necessary to classify their components in different ways to reflect the point of view that is relevant for different kinds of analyses.

In addition to the subdivision of *transactions* it has been necessary to provide for the subdivision and alternative forms of classification of the broad categories of *transactors* which have been identified in the summary accounts.

Thus we would look to the domestic production account being divided up into accounts for each industry and for the consumption (income and outlay) accounts to be subdivided for different sectors and sub-sectors of the economy.

In doing so, many more transaction flows will need to be made visible. For example, *direct taxes* and *transfers from government* will appear as both receipts and payments when the consumption (income and outlay) account is subdivided to show government separately from households. These represent flows between units ***within*** the consolidated income and outlay account for the nation as a whole, where they would normally be netted out.

Real and financial perspectives

two preoccupations—

(1) production and income generation

(2) distribution of the income and financing the process

In the continuing interaction between the designers and developers of national accounting systems and the economic analysts using the accounts and supporting statistics, there have been two broad preoccupations. One is with the process of production and income formation, while the other is with the distribution of this income and the process of financing the production process.

Necessarily these are sub-processes of an overall process of economic growth and change — the production process cannot exist without the financial processes and vice versa. Nevertheless the two fields are distinctive enough to have been treated separately in the development of statistical support systems like input-output tables, on the one hand, and income distribution, flow of funds and national balance sheet work on the other.

The 1968 SNA sought to contain and relate both of these broad fields of interest within its accounting design. It provided a comprehensive framework for

(a) integrating production accounts with input-output studies, and

(b) integrating income and outlay data with capital finance accounts, flow of funds and balance sheets.

With country experience in this, SNA Revision 4 is now directing attention towards ways of representing more plainly the linkages between (a) and (b).

two perspectives

Essentially, the distinction between the real and financial perspectives turns on the fact that the former is concerned with physical/technological transactions, as far as real flows can be represented in money value, while the latter focuses on money value transactions only.

From the *real* perspective, one is trying to observe, as closely as possible, the physical process of production and the process of price and income formation as it takes place via the market.

From the *financial* perspective, one is trying to observe how money incomes flow to owners/consumers, how they are redistributed and how they are used to purchase the final products of the production process and to finance further production.

Two ways of dividing the economy into sectors

The designers of the SNA recognised the need to define and classify different units of observation for the purpose of analysing transactions from the two different perspectives.

institutional units of the enterprise type make the fundamental decisions

The SNA first recognises units of the enterprise-type which it calls *institutional units*. These are the fundamental decision-making units in society, which can engage in the full range of economic transactions and which are legally responsible for their activities (eg they are capable of owning assets and incurring liabilities on their own behalf).

They are grouped into sectors and subsectors so that their distinctive roles in the economic system can be analysed in a full range of accounts.

The broad sectoring of institutional units proposed for SNA Rev. 4 is into —

(i) *financial corporations* (such as banks),

(ii) *non-financial corporations* (the bulk of producers of goods and services)

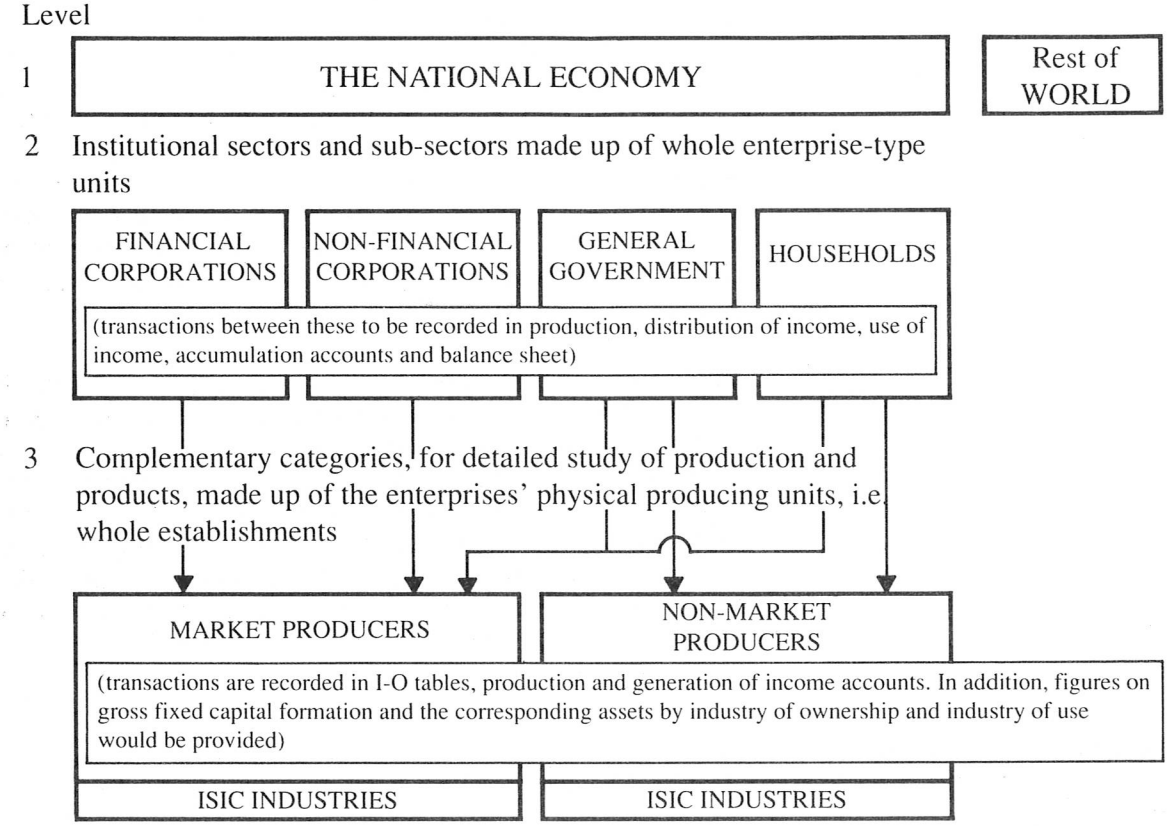
(iii) *general government* (which in their economic role produce principally non-market services for individual or collective consumption and redistribute national income and wealth) and

(iv) *households,* which may be both producers and consumers.

The household and the unincorporated enterprises of a household (eg the corner shop which is also the household's home) are included together as one owning/financing unit in the Household Sector. The sector also includes non-profit institutions serving households (eg community agencies providing non-market services for households, supported by voluntary contributions or grants).

As the next diagram indicates, to analyse the process of income distribution and financing, the SNA prescribes income and outlay, capital finance accounts and balance sheets. The unit of observation for this is of the enterprise-type — broadly the units in which financing decisions are taken.

PROPOSED SNA REV 4 SECTORS



NOTE: ISIC is the International Standard Industrial Classification

establishment-type
physical
producing
units

Because the enterprise-type units can include large corporations, with many different types of productive activity at different locations, the SNA provides for more detailed analysis of the process of production and income formation by identifying producing units which are as homogeneous as possible (ideally engaged in only one activity and visible at a particular location). They could be the only physical establishments of an enterprise or one of a number of producing units owned by an enterprise.

Detailed production accounts for industries and commodities are prepared from data obtained about the characteristics and transactions of these types of units — "*broadly, the smallest organisational units in which decisions are taken on production matters and for which basic production data might be made available.*"

These establishments are classified according to their principal kind of activity and then grouped into industries. They are classified into broad functional sectors distinguishing between those mainly producing for the market and those which are not. Additionally, there would be production accounts for market goods and services to be analysed as to their source and use.

The use of the two different levels of statistical unit, grouped in alternative functional categories, also recognises the fact that complex enterprises (which may consist of a number of corporations each with a number of component operating establishments) may record financial transactions, such as capital raising, only for the enterprise as a whole. **We therefore cannot work with as fine a level of unit of observation for analysing financial transactions of businesses as we can for analysing their production transactions.**

Note that concepts and terminology vary in the way in which different countries define institutional (owning) units as distinct from establishment type (producing) units. For example, some use the term *enterprise* to refer to the (operating) legal entity while others use it to refer to the enterprise group. Some have put the main emphasis on the establishment being a physical location, for which production data are required; others have made *availability of data* the key criterion of the existence of an establishment.

The latest Australian approach is to recognise, for statistical reporting, the enterprise and the enterprise group and three levels of establishment type unit, the management unit, the establishment and the location. Thus

management unit — generally the highest unit within an enterprise group hierarchy for which at least quarterly accounts of income, expenses, employment, capital expenditure and stocks are maintained;

establishment — the smallest unit within a management unit which controls its productive activities and for which a specified range of detailed annual data on income, operating, employment and stock data are maintained;

location — relatively permanent physical sites occupied by an establishment and for which employment data are collected.

Input-output analysis

relates to flows of goods and services; production and income generation by industries and establishment-type units

While it is not practicable in this *guided tour* of national statistical systems to pursue this topic in any depth, its place in the SNA statistical unit structure needs to be indicated at this point.

Input-output studies aim to provide detailed information, presented in matrix form, about the supply and disposition of goods and services in the economy and the structure and inter-relationship of industries.

They could be viewed as being, in principle, a disaggregation of the SNA Rev. 3's production accounts which are of two types, showing —

(1) the *supply and disposition* of *commodities** and of *other goods and services* (flows of goods and services by class of product) and

(2) *production* by *industries* (ie production and generation of income of groups of establishments classified by kind of activity).

* *Note that SNA Rev 4 is dropping its specialised terms 'commodities' and 'other goods and services' in favour of 'market products' and 'non-market products'*

The SNA production accounts are primarily concerned with presenting the end results of economic activity — the final value-added net product from which income and final consumption arise.

The input-output (I-O) tables, by explicitly showing the intermediate transactions, complete the description of productive activity within the economy and provide the starting point for many analyses which could not otherwise be undertaken.

In some countries, the I-O tables are designed to be an integral part of the build up of these production accounts and, in turn, of the consolidated national production account (in Chapter 14 we will illustrate how input-output tables are logically part of the national accounting process).

However, in many countries, the I-O tables remain separate exercises, undertaken periodically, and the results obtained are not identical with the national accounts because of differences in estimation methods, classifications and data sources. For example, industries may be redefined from those specified in the national industry classification in order to better analyse physical processes, such as energy consumption and the effects on the environment.

In the ideal I-O world, industries could be defined as coinciding with product — eg the ball bearing industry would consist of establishments which make ball bearings as their primary (and only) product and there would be no other industries also making ball-bearings as a secondary product.

In fact, however, industries are made up of physical establishments for the purpose of identification and data collection in censuses and surveys and these establishments do have mixed activities (eg roller bearings may be produced in ball bearing establishments).

So industries are defined so as to be made up of establishments which are as *homogeneous* (similar in nature) as possible in terms of the commodities they use up in production.

This is a problem for the design of industry classifications as well as for constructing I-O tables on a basis which is both efficient for the applications to which I-O tables are put and for integrating national accounts and basic industrial statistics

In fact, the I-O solution sometimes is to shift some secondary production, together with the relevant inputs, from the industry where it is produced to the industry to which it is primary.

Its industries then consist of more homogeneous units of production which may not be physically observable establishments as defined for economic censuses and economic accounts. This will accommodate the technological point of view but at the expense of a faithful description of the real world of economic units and their activities as they are actually organised, managed and accounted for.

Note also that, while the SNA tends to reserve the term 'sector' to institutional sectors, sector in I-O usage can refer to classes of goods as well as to classes of industries, with the inputs and outputs between them recorded in the rows and columns of either a commodity by commodity matrix or an industry by industry matrix.

The term 'commodity' as used above refers to the products characteristic of an industry and is not, in this usage, limited to products produced for the market (as SNA Rev.4 uses the term).

Flow of funds

focus on financial transactions between institutional units — classified by type of instrument and sector

Whereas I-O studies analyse the economy at the establishment and industry level and in terms of commodity flows, flow of funds analysis relates to *enterprise-type units*, classified to *sectors* and *subsectors,* and in terms of *financial flows analysed by type of instrument.*

This field of economic accounting is now well integrated into the SNA's financial accounts and should no longer raise problems of inconsistency in respect of definitions and classifications of transactors.

The main purpose of flow of funds accounts is to provide a framework for analysing the interrelationships between savings and investment decisions in the *real* economy (eg in producing, accumulating and using up capital goods), and the underlying financing activity.

Thus they can be viewed as concentrating attention at a sectoral level on the transactions summarised in the SNA's capital accounts (dealing with the accumulation of assets) and financial transactions accounts.

The starting point for this is *saving* carried over as the last balancing item from the *current accounts* of the system (which deal with production, income and use of income)

Saving then is shown on the right-hand (resources) side of the *capital account* along with *capital transfers* receivable. The balancing item on the uses side of this account is *net lending* and should in principle be identical with the *change in financial position* which is the balancing item of the *financial transactions account.* Operationally there would be discrepancies between the two items in that they are compiled from different data sources.

Note that the change in financial position can be built up independently by sector from data on transactions in financial instruments in categories significant for monetary analysis.

This is illustrated in the following diagram from an Australian publication *Australian National Accounts Flow of Funds Developmental Estimates* (ABS Catalogue No. 5232.0)

NATIONAL ACCOUNTING RELATIONSHIPS

Income and Outlay Account

Final consumption expenditure	Wages, salaries and supplements
Current transfers paid	Net operating surplus of trading enterprises
Saving	Current transfers received
Total disbursements	Total receipts

Capital Account

Fixed capital expenditure	*Saving*
Increase in stocks	Consumption of fixed capital
Net purchases of land and intangible assets	Capital transfers received
Net lending	
Gross accumulation	Finance of gross accumulation

Financial Transaction Account

Cash and deposits	Cash and deposits
Loans	Loans
Short term securities	Short term securities
Long term securities	Long term securities
Equities	Equities
Other claims	Other claims
Change in financial position	
Net incurrence of liabilities	Net acquisition of financial assets

analysis by sector and type of financial instrument (the commodities of financial services)

For the purpose of providing more detailed analysis of transactions in financial claims, institutional sectors and sub-sectors made up of enterprise and household type units can be further divided up into homogeneous classes of institutions which each have similar roles in income distribution and redistribution and in the capital financing process.

This is done with a concern for the logical links with the production process and other related systems — notably balance of payments, money and banking and national balance sheets.

We will touch on this field again in Chapter 16 as we focus on the characteristics of enterprise and higher level data collections and the special purpose derived systems which they support.

Domestic territory and resident economic agents

Before we switch attention to the Rest of the World Accounts we need to define the boundary between the domestic economy and the rest of the world.

The conventions relating to the concepts of territory and residence are central to national accounting and flow through to economic statistics generally.

why gross domestic product?

The *domestic territory* of a country is defined to exclude the overseas territories and possessions and to include, in addition to the territory lying within its political frontiers —

(i) ships and aircraft which resident enterprises of the country run entirely, or primarily, between two or more countries,

(ii) fishing fleets, vessels and floating platforms which residents of the country operate wholly, or mainly, in international waters, and

(iii) fishing vessels, oil and natural gas rigs, and platforms engaged in extraction in areas in which the country has the exclusive rights of exploitation by virtue of international agreements or pronouncements.

The general principle adopted for undertakings which operate across borders is that the production, and the income arising from this, should be ascribed to the country in which the production takes place and resident producer units are to be delineated for this purpose.

production — a territorial concept

The accounts in respect of the production of goods and services cover the transactions — for example, gross output, intermediate consumption, value added — of all resident producers.

Resident producers are defined so that all (and solely) production taking place on the domestic territory of the given country is encompassed. **The production accounts of the system therefore relate to the activities of resident producers, and not to the return to resident factors of production.** In other words, the territorial concept of domestic production is used.

final consumption expenditure of resident economic agents

The accounts on the disposition of goods and services are also to portray the transactions of resident economic agents. The final consumption accounts cover the outlays

on this in respect of resident households, producers of private non-profit services to households, and producers of government services.

Again, the capital formation accounts deal with the activities of resident producers only. Fixed capital formation consists of the acquisition of fixed assets by resident producers and occurs only on the domestic territory of the given country.

Additions to stocks also represent capital formation. In principle they are included as the capital formation of resident economic agents as soon as ownership has passed to them from non-residents, whether or not the goods have crossed the frontier. In practice their importation will not be recorded before they have passed through Customs.

income is a residential concept

Resident economic agents receive compensation as employees and income from property and entrepreneurship as a result of participation in the production of the given country and in the production of the rest of the world. Some of the incomes accruing from the production of the country in question are also paid to non-resident economic agents.

Thus the incomes accruing to resident economic agents for contributing to production both inside and outside the domestic territory differ from the incomes accruing from domestic production.

centre of interest

The concept of residence adopted in the case of individuals and households is designed to encompass all individuals who may be expected to consume goods and services, participate in production and engage in other economic activities in the domestic territory of a country on a lasting basis (one year or more is taken as a practical guideline). These are the persons whose general centre of interest is considered to rest in the given country.

Rest of world accounts and balance of payments

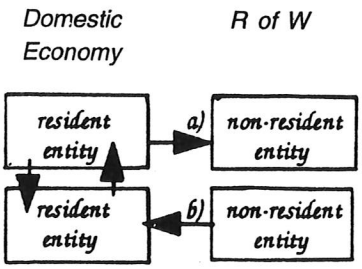
We have seen that the transactors and holders of stocks under the national accounts system are the resident economic entities of an economy.

The SNA analyses their activities in terms of three broad categories of accounts:

- *current accounts,* relating to production, distribution of income, and use of income;

- *accumulation accounts:* relating to changes in assets and liabilities and changes in net worth;

- *balance sheets*: relating to stocks of assets and liabilities and net worth.



In order for the system to be a *closed* one (ie one which can be balanced) a segment must be provided to which a) outgoings, from one resident entity that are not incomings for another resident entity, may be directed and in which b) incomings to one resident entity, that are not outgoings for another resident entity, may originate.

That segment is known in the SNA as the *rest of the world* (ROW). The segment occupied by the resident entities is divided among a number of sectors with accounts of production, consumption and accumulation, as we have discussed.

monitoring external transactions

The broad interest in monitoring external transactions has to do with transactions on current and capital account, covering flows between residents and foreigners of goods, services, income and other transfers, and investment. These transactions are the subject of the *Rest of the World (or External Transactions) Account and Balance of Payments statistics.* Detailed commodity breakdowns are provided by *International Trade Statistics (import and export statistics).*

— on current account

The commodity flows monitored in the International Trade Statistics are, together with other goods, services, income and transfers data, inputs to the ROW current account and Balance of Payments current account.

— on capital account

The counterparts of these flows are the flows of investment (including changes in the international debt position) which are inputs to the *Balance of Payments capital account* and *ROW capital account* and in turn to statements of the *International Investment Position.*

Also, the flow detail on goods and services is an input into the SNA production accounts and into the related, more detailed input-output analysis of the production process.

Historically, the *Balance of Payments* is a specialised information system developed by the International Monetary Fund (IMF) for the purpose of supporting its activities in promoting orderly international exchange

relationship of B of P and SNA

Basically the IMF uses the same building blocks of statistical units and data items and applies the same kind of accounting processes as does the SNA.

Thus the differences between the SNA and Balance of Payments system have been more differences in presentation than differences of real substance and changes are now being made to bring the two into line as part of the economic accounting system directly linked with the SNA's central system.

The main purpose of the SNA is to exhibit the interrelationships between the rest of the world and the total national economic system. Thus it can be tapped readily into a variety of economic analyses in such fields as overall economic planning and policy formulation, input-output, production functions, factor shares, consumer demand, distribution and redistribution of income, and flow of funds.

The Balance of Payments presentations are a more specialised tool for identifying imbalances between the domestic economy and the rest of the world and for tracing the effects of those imbalances and their financing on international relationships and the adjustment process. They include additional information besides that which appears in the SNA, such as total change in reserves and other balances that are of special interest in analysing external transactions while not fitting directly into the SNA classifications.

In fact it is normal practice for Balance of Payments items to be the main source for the national accounts and for flow of funds compilation.

Thus the flow of funds accounts, in the rest of the world sector, will contain the capital account of the Balance of Payments.

presentational differences

The different perspectives of the Rest of the World and Balance of Payments presentations sometimes confuse the uninitiated.

In the Balance of Payments, as in business accounting, *increases in assets* (eg imports of real resources, lending abroad) and *decrease in liabilities* (eg repayment of debt) are recorded as *debits* (minus sign) and *decreases in assets* (eg exports of real resources, a sale of foreign securities) and *increase in liabilities* (eg the utilisation of foreign loans) as *credits*.

While the SNA is a closed system and records two ends to every transaction, the balance of payments is an open system which shows only the resident end of every transaction. To grasp the distinction more clearly, remember that each economic transaction may be regarded as two flows.

For example, in the case of an import which is paid immediately by cash we can record

	Resident	*Non-Resident*
Real flow	*1) debit*	*2) credit*
Financial flows	*3) credit*	*4) debit*

The Balance of Payments (B of P) is an *open system* dealing only with the resident aspects of the resident and non-resident transactions and records only entries for 1) and 3).

However, the SNA is a *closed system* recording the transactions for all the parties to transactions, including non-residents — in this case it would record entries for 1) and 2) from the point of view of the rest of the world in the Rest of the World account and from the point of view of residents in the domestic production account.

Entries 3) and 4) from the point of view of residents would be shown in the flow of funds account.

The B of P estimates are compiled from a wide variety of administrative sources and special surveys. The major source for covering imports and exports of goods in the balance of payments is what is referred to as *international trade statistics* which are derived from the administrative records of Customs agencies.

the B of P is not just about money transactions and balances

The term *Balance of Payments* hardly does justice to the nature and scope of this statistical field. The B of P summarises systematically all the economic transactions between residents of one country and residents of the rest of the world —

- not just payments, not just cash holdings
- all we owe and are owed
- all flows of real resources (goods, factor and non-factor services)
- all transactions in financial claims between residents and non-residents

The summary is backed up by a body of systematically organised data for analysis

- on trade in goods and associated prices

- *invisibles* i.e. services (eg freight and insurance) and income transfers

- private and government financial transactions

- balance sheet information on (eg) foreign exchange holdings and position with the IMF.

Thus, a study of a typical balance of payments presentation will show that it is split into a current account and a capital account with the balance on current account carrying into the capital account and with other balances of interest shown.

Adaptations to country circumstances and needs

While most UN member countries accept the SNA and its various extensions as the world standard, there are many variations on its theme as it has been adopted by different countries and by international and regional authorities (such as OECD and EEC). The introduction to its design principles in the previous chapters will generally provide useful leads as to the rationale behind what has actually been done in each country in accordance with their particular circumstances and needs.

The UN guidelines and country systems are constantly being elaborated as experience and aspirations increase. A framework of related concepts in a structure of balanced accounts helps to ensure that coherence is maintained. The SNA's general accounting scheme forces consistency between items, both as a theoretical system of ideas and as a system for presenting collected and estimated data. As the next 2 chapters will demonstrate, an understanding of its principles is basic to an understanding of economic statistics generally.

CHAPTER 13

EXTENSIONS BEYOND THE CENTRAL FRAMEWORK

We have so far concentrated on the what and why of the central framework of the SNA and its elaborations. Our purpose in this chapter is to move on from the relatively well ordered fields of activity which can be analysed in a fully articulated system of accounts to explore the substantial marginal lands which can potentially be viewed in terms of money values but with much greater difficulty or uncertainty.

New pressures from changes in social priorities

The 1968 SNA revision concentrated on the internal consistency aspect of national accounting and the integration of economic statistics rather than on the validity of the traditional aggregate measures of economic progress. Statisticians then responded magnificently to its basic challenge to provide a comprehensive and detailed framework for recording the stocks and flows of the economy.

But this has been overtaken by a new demand for measures of progress introducing concerns which are not readily accountable in terms of economic transactions.

In particular, it was seen that no adequate account was being taken of the deterioration of the natural environment, the depletion of non-reproducible natural resources and the disamenities that arise in a society geared to maximising income in the conventionally limited transaction-accounting sense.

shortcomings of GDP as a measure of progress

While some aspects of these concerns will remain beyond the scope of a substantially transaction-based accounting system, some are within scope in principle and are being addressed in the near-completed current review of the SNA and in associated experimental statistical research and development.

Specifically —

Distribution of income and wealth — it was evident that, however estimation might be refined, a measure of the total amount of income and wealth in a nation would not be a satisfactory indicator of welfare if it did not take account of inequities in the distribution of that income and wealth.

It was recognised that, although the 1968 SNA was intended to provide a framework for analysing the distribution of income and wealth as well as national aggregates, it had failed to stimulate much progress in this field.

production which has worth but is not paid for

Omission of elements capable of money valuation — while some imputations have always been made in the SNA for non-market agricultural production, the services of owner-occupied housing and services provided by financial intermediaries, there is an argument for estimating the value of unpaid household work and *do-it-yourself* activities and to include intangible and human capital as well as tangible capital. It has also been argued that the treatment of owner-occupied dwellings as capital equipment, providing a flow of services to the owner, should be extended to other household durable equipment and to government capital equipment.

production which is paid for but has no positive worth

Inclusion of negative aspects of production — by the same token there are other imputations or adjustments that might be made to *reduce* GDP by adjusting it downwards for transactions which do not contribute to quality of life or which detract from it.

Thus it can be argued that some activities which have been treated as final consumption (eg *defensive* or *regrettable* expenditure by government and industry in maintaining garbage and pollution abatement services or by households in commuting to work) should be treated as intermediate consumption — a cost to be deducted in deriving a net social product.

Again, there is the need to bring to account the negative effects of economic development on the natural and built environment and on the stock of both renewable and non-reproducible natural resources needed to sustain development.

What should be apparent is that these are areas of imputation for which it is difficult to be objective and consistent. The values that we might place on such activities and effects are neither stable nor consistent between cultures in this respect.

We do not have to go far into the conceptual and practical problems implicit in extending money valuation beyond the solid ground of actual transactions to realise that international consensus on standards will not occur overnight.

How do we value aesthetic and other environmental considerations which are of high concern to those in a position to enjoy natural amenity but which are not priorities for (eg) hill tribes whose livelihood depends on 'slash and burn" cultivation? And how to value extra-market activities which are productive of household goods or services but are undertaken largely for personal enjoyment and at costs which generally could not be justified under market conditions — eg, our 'do it yourself' craft activities and the time we spend with our own children is generally much more than we would be prepared to pay for.

Or how do we put a value on such qualities of life as security? And should we regard government spending on defence or on internal order as a measure of economic growth? And how do we cope with valuing the different qualities of political and religious freedom evident in different countries at different times?

Nor would we have to go far in this process before our imputations will far exceed market-based valuations as, in terms of the diagram at the end of Chapter 10, we head towards the *outer space* of non-economic valuation where attempts to stretch the national accounts are certainly no longer feasible or meaningful.

what imputations are worth attempting?

Nevertheless, beyond the 'core' national accounts is much that is capable of market valuation with some difficulty and is significant enough for social analysis purposes to warrant the attempt — if only to put the key national aggregates into perspective.

In fact, while it is fashionable to write off national accounting as an instrument for social analysis, the detractors really offer no alternative to a money-value accounting approach when they argue the need for more inclusive measures of economic welfare. *Given the state of the art, how far is it really worthwhile to pursue this course?*

The answer to this question depends very much on how we rate the importance for social analysis of the unique attributes of a money-value accounting system. Lets look at these attributes again.

What are the systems dimensions of the accounting approach?

In considering the products of national statistical systems we chose to begin with the economic accounting aspect of social statistics because it forms a centre-piece, with something to contribute to every field of social statistics, and because it seems to be the most developed model for the systematic organisation of statistics.

Certainly it is worth identifying the characteristics of this very comprehensive and coherent system in the light of the aspirations of those who would like to extend it across (or to link it with) other fields of social enquiry.

the common money measure and arithmetic identities

The use of the common measuring rod of money valuation gives accounting and economic accounting an *adding up* quality which the SNA exploits to the full in its arithmetic identities relating production, income, consumption, saving, financing and accumulating, and in its connecting of stocks and flows.

As we will see in the next chapter, the consistency and relatedness of its concepts and classifications for the actors and activities covered can be checked operationally at many points through the confrontation and necessary reconciliation of the constituent sets of data from different sources.

consistent with — a general theory of social behaviour and with commercial accounting

Moreover the system is based on well established theory as to how the economy works and on concepts and data requirements which are substantially consistent with the conventions and practices of record keeping at the micro-level of the individual economic units (and in turn, with primary data collection systems based on these).

Consequently, SNA-based economic statistics represent a very extensive, open-ended and flexible framework for analysis at many alternative levels of aggregation and in respect of many alternative categories of actors and activities — not just national aggregates.

translation to physical units

It should be recalled also that, while money value recording of flows and stocks is necessary to balance up the various accounts, it does not prevent behavioural models (and, in turn, our data systems) from expressing the fact that at all levels we monitor and control many activities in physical unit terms (though often within a broader financial framework of budget limits).

Thus, to recapitulate Chapter 10, there is a natural connection between money value data and other data by way of the price (or imputed price) of each quantity which we regard as having a value in use or exchange — ie P x Q = V.

Of course, when we denominate statistical series in such physical units as **cannot** be converted to money value data, we do without the system potential of a money-value accounting framework. Key summary indicators like infant mortality and literacy, recorded as numbers of persons, would be in this category.

Extensions and elaborations of the national accounts for broader social purposes

New Economic Welfare Measures

Some pioneering explorations of these possibilities have demonstrated how far money-value accounts might usefully be adjusted and extended to provide better measures of total economic welfare.

For example, Nordhaus and Tobin attempted a measure of economic welfare which 1) excluded intermediate or instrumental expenditures, 2) included the value of capital services of consumer durables and government capital and the value of leisure and non-market work, and 3) made deductions for the disamenities of urbanization. Thus they argued that consumer expenditures for commuting to work are instrumental rather than final consumption, and that many government purchases were either 'regrettables' (such as national defense) or intermediate services (such as sanitation and police).

'Is growth obsolete' — National Bureau of Economic Research, Fiftieth Anniversary Colloquium V, Economic Growth (Columbia University Press for National Bureau of Economic Research, New York, 1972).

Such studies did not have much impact on official statistics — rather they have highlighted the enormity of the conceptual and data collection problems involved. But

with research experience and discussion in academic forums and with follow-up development by national statistical services they can be expected to progressively gain legitimacy.

SNA Revision 4's horizons

The current review of the SNA gives a fair picture of the extent to which national accountants are looking beyond the conventional production boundary or are trying to provide alternative breakdowns. Social analysts should be aware of what is being recommended for the forthcoming Revision 4.

See (eg) UNSO provisional future ST/ESA/STAT/SER. F/2/Rev.4 — Revised System of National Accounts; Preliminary Draft Chapters Chapter II: An Overview of the System, Feb 1990) or Anne Harrison, Major Changes Proposed for the Next SNA: An Overview, Review of Income and Wealth, Series 36, No. 4, Dec.1990.

beyond the boundary

While the *boundary of production* and, consequently GDP and related aggregates, remain substantially as before, there are a number of ways in which national accounts may serve such social purposes as we have identified. Thus —

The system provides for much more detailed analysis of the household sector with socio-economic sub-sectors and further disaggregations carried through into the accounts and tables. The elaborated accounting structure also pays great attention to income distribution and redistribution, including redistribution in kind.

Its detailed analysis of household and government expenditure by purpose will also be of special interest for social analysis as it identifies the type of need a transaction aims to satisfy or the kind of objective it pursues.

Social Accounting Matrices

These developments, along with the adoption of extra breakdowns to suit individual country needs (eg into urban/rural households, by industry of employment, education, skills etc) open up a whole range of analytical possibilities (such as the distributional effects of income generation and distribution) using *social accounting matrices (SAMS).*

SAMS are a comparatively recent development now being recognised formally as a natural and valuable disaggregation and restructuring of the SNA. They show how sub-sectoring as envisaged by the SNA can provide for the useful incorporation of social and demographic information within the SNA's existing framework so as to use the SNA's natural flexibility.

SAMS portray those macro-economic links and financial inter-relationships within a country which are important for economic and social development strategies working through the distribution of incomes and the composition of household demand.

Thus, both macro-economic aggregates (the growth indicators) and distribution and redistribution (through taxes etc) can be recorded in a matrix presentation and so integrated. (see Keuning and Ruijter, Guidelines for the Construction of a Social Accounting Matrix, Review of Income and Wealth, March 1988)

Development of Macro-micro Analysis

The possibilities of *macro-micro links* are introduced in the SNA review as another dimension along which analysis may be deepened within the central framework.

The system as it stands can be described as a macro/meso economic framework in which global figures are subdivided by sectors and subsectors with cross classifications by industry, product, region etc to provide a substantial data base for socio-economic analysis, model building and so forth. **But, in the aggregation and estimation process, much information of interest disappears behind the average figures, even at a highly disaggregated meso level.**

An example, given in preliminary papers on the revised SNA, is that interpreting the global household saving ratio may necessitate taking into consideration the age and socio-economic structure of the population, its rural/urban distribution etc.

Computer facilities, data bases with sequences of data from statistical units (households and enterprises) open up possibilities of linking the micro data with the macro/meso figures and vastly expanding the system's interpretative power and flexibility.

Richard Ruggles sees great possibilities in utilizing longitudinal survey or administrative data for micro-simulation modelling related to macro-economic analysis. In the United States' context, he suggests that —

'a relatively modest household microdata set that is integrated with (ie., consolidates to) the household sector of the national accounts could yield useful disaggregations of the major items of income and expenditure, and provide related social and demographic information'. (Survey of Current Business Nov.1982, p41)

Given the advances in computerization, the more difficult problems of achieving macro-micro integration are likely to lie particularly in the area of accounting standards and (for households) in the conceptual and operational co-ordination of survey and administrative by-product systems.

Satellite accounts

Extensions beyond the central framework are envisaged by way of *satellite accounts* in order to

(i) cover more completely fields or aspects which are not fully developed in the central framework, including the introduction of complementary or alternative concepts, etc or

(ii) focus on a certain field or aspect of social life, in the context of national accounts but without at the same time covering the rest of the economy.

Thus, it is suggested that important fields of economic and social concern, such as environmental protection, education and health, may benefit from building up a framework which might contain both —

(a) elements explicitly or implicitly included in the central accounts, and

(b) complimentary elements (in either monetary unit or physical quantities) and even alternative concepts, definitions, classifications and presentations that better serve specific purposes in the particular field.

To illustrate this approach, it is observed that such a set of satellite accounts '.....would allow specific areas of social concern to be analysed, without overburdening or disrupting the 'core' system. The satellite accounts would provide measures of defensive expenditures and of the use, depletion and degradation of natural resources. This would enable aggregate indicators of environmentally adjusted and/or sustainable income and product to be produced. The satellite accounts would also link physical resources with monetary environmental measures and balance sheets. The view of the national accounts experts is that further consideration could be given to adjusting the core of the SNA once adequate experience has been gained with these types of accounts and when various conceptual and valuation issues have been resolved.'

(Feature article: Natural Resource and Environmental Accounting in the National Accounts, Australian National Accounts, National Income and Expenditure, March quarter 1990 — ABS catalogue 5206.0).

A change of perspective

Enough has been said on economic accounting design to indicate that national accounting, with its natural extensions and elaborations, is strengthening its role in social analysis and should be encouraged to do so.

But, notwithstanding the power of an accounting system for monitoring and controlling our activities, decisions in society are not made purely on the basis of priorities expressed in actual or imputed market values. In Part 4 we will move out of the economic accounting context to consider systems which use other dimensions than market values, but which will still need to maintain substantial links with economic accounting systems and analysis.

In the remainder of Part 3 we shift attention from objectives and design to consider operational processes and the nature of the data base for economic statistics.

CHAPTER 14

THE ECONOMIC ACCOUNTING PROCESS

We have viewed economic accounting systems as a framework of concepts and classifications of transactions and transactors within which can be organised information about society's activities in producing and consuming goods and services. This chapter now looks at the processes by which the specified accounting elements are assembled from the many and various sources of information.

Filling the 'empty boxes' of the economic accounting system is no mean task. After all, it involves bringing to account every monetary transaction of every economic agent in the economy, plus the imputed transactions of producers who consume their own product. Since we know that only a fraction of such transactions is reported to government statisticians, how are national aggregates like GDP (and all the accounts and tables that relate to the key aggregates) produced regularly at annual and perhaps quarterly intervals?

Alternative accounting routes

The discussion in chapter 11 gave a clue to the accounting strategy as it focused on the alternative ways of viewing the product of an economy. Production, income and expenditure are equivalents (at least in a Robinson Crusoe type economy or for the world as a whole).

GDP is the sum of the net output of producers (plus import duties), or of incomes arising from production, or of outlays on the disposal of production in different uses. **Corresponding to these three different views of the composition of GDP there are three alternative approaches to measuring it.**

In practice the statistician will make the best of the data available on each of the three routes to GDP.

Ideally the statistician can monitor the accounts of producers directly to collect information as to their value added. This is the *production approach* or *net output method,* represented by the output less the input — the contribution by each producer to GDP.

How this could be derived at micro level can be illustrated from the following simplified business production account —

Sales		*$15,000*
Increase in stocks:		
finished goods and work in progress		*1,000*
		$16,000
Materials and fuel purchased	*4,750*	
***less** increase in stocks of materials and fuel*	*250*	
Intermediate consumption	*4,500*	
Wages	*6,000*	
Indirect taxes	*1,000*	
Depreciation	*500*	
		12,000
BALANCE — Operating surplus		*4,000*
		$16,000

The total gross output of goods produced in this example is sales ($15000) plus increase in stocks of finished goods and work in progress ($1,000), that is, $16,000.

production approach

If the total value of output for every producer were added up across the whole economy we would arrive at an estimate of the nation's gross output. However, this would not be a meaningful concept because of the double counting of the goods and services consumed by each producer in producing their total output for the period.

Hence, the intermediate consumption ($4500) of the producer is deducted from the value of output ($16,000) to determine the value added by the producer ($11,500).

The aggregate of the value added by all producers, gives GDP by the production approach (ie in terms of net output).

income (cost) approach

The lower half of the account sets out the producer's costs, including the balancing item *operating surplus*, which is available to pay those who have invested in the business.

With the exception of the materials and fuel purchased (intermediate consumption), these are the items required to measure GDP by what is referred to as the *income approach* (or *incomes distributed* method), since they represent the incomes arising from production, plus the indirect taxes accruing to government and also depreciation (GDP being gross of depreciation).

From the business point of view, these are all *costs* and the UN prefers to refer to this as a *cost approach* in its *Handbook of National Accounting* (Studies in Methods, Series F no 39)

Thus, adding all these items, except the $4,500 intermediate consumption, again gives a GDP contribution of $11,500.

When elements of this are obtained directly from producing establishments, along with the items required for the net output method, it could be said that a production *approach* was being followed to provide GDP estimated by both the net output and the incomes distributed *methods.*

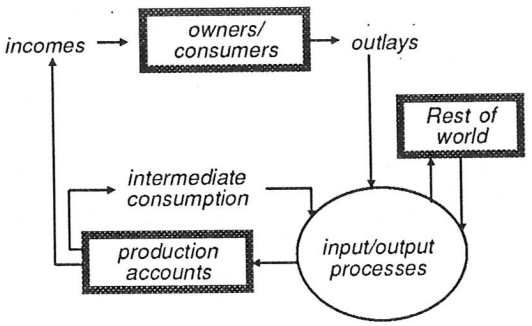
expenditure approach

The third approach is referred to as the *expenditure approach* **in the UN Handbook of National Accounts. In the literature it is also variously referred to as the** *consumption approach* **or the** *incomes disposed* **or** *final products* **method.**

The expenditure approach involves monitoring the goods and services being exported during the period or being finally consumed or added to the domestic economy's productive capacity as fixed capital formation or as net additions to stocks. These are the final outputs from which the imports are to be deducted (see diagram '3 Faces of GDP ' — Chapter 11). Since this way of visualising and measuring GDP draws attention to the physical products emerging at the end of the production pipeline, the term 'final products' method has some appeal.

Economic transaction flows — points of observation

If we return to a simplified macro model of the economy we can visualize the three methods in terms of different points of observation for monitoring the flow of money value transactions.

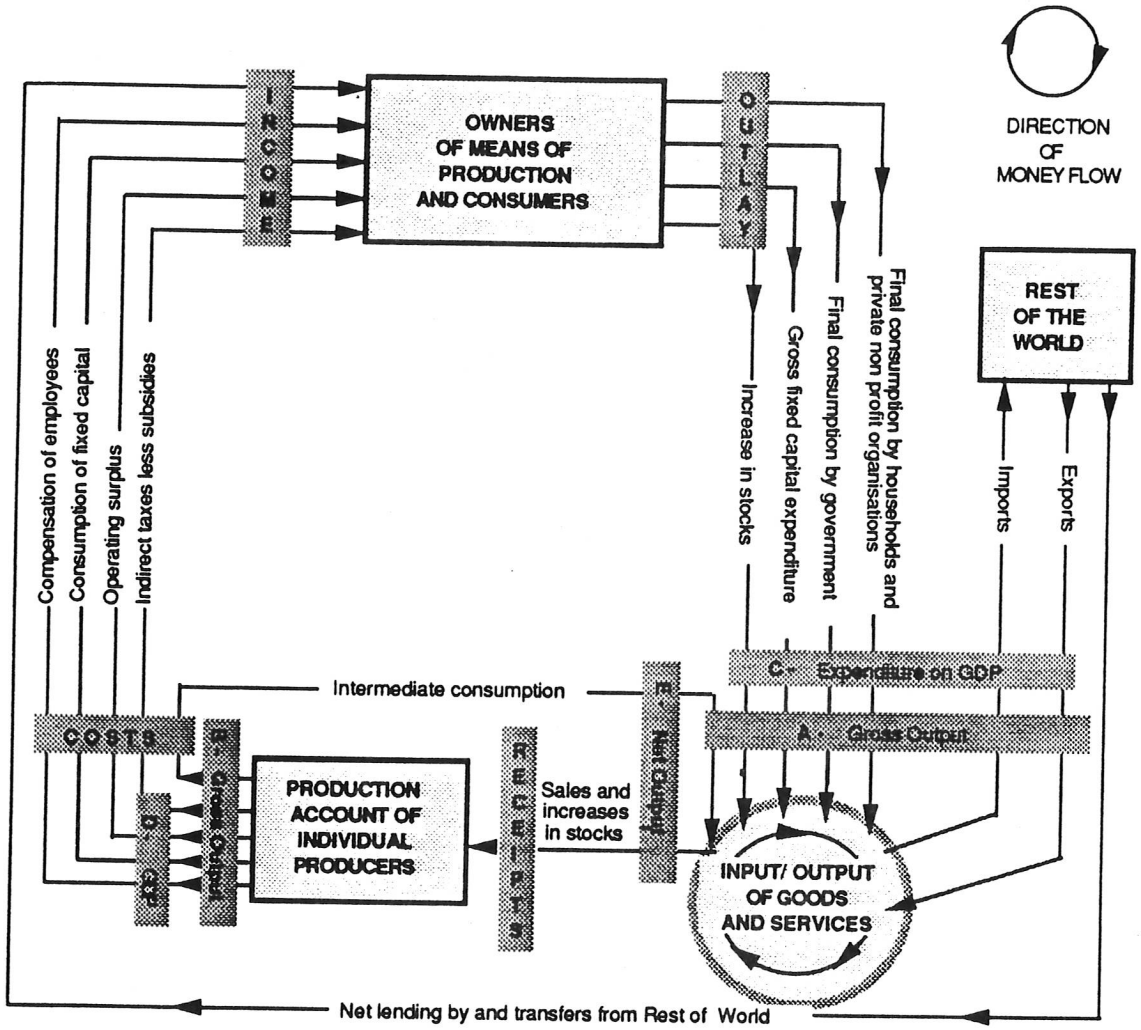


On the left of the diagram (elaborated over page), returns to the factors of production are flowing to owners and consumers from producers. This is the distribution of the proceeds of production.

On the right of the top box are shown payments for final goods and services by the owners/consumers.

Together with payments received from the Rest of the World for exports of goods and services, and after imports have been paid for, the balance flows as payment to the owners of the means of production.

TAPPING TRANSACTION FLOWS FOR ALTERNATIVE MEASURES OF GDP



This pictures the outlays on (eg) gross fixed capital formation in the domestic territory as purchases from domestic producers and from the rest of the world. The flows between different sources and uses of the different kinds of goods and services (including intermediate goods and services) is represented as taking place within the circle labelled input/output of goods and services. These flows of commodities between industries are analysed in interindustry studies, whereas the standard national accounts focus on the final outcomes of this process.

This rather daunting boiler diagram is not intended to be memorised! It is simply intended to highlight the fact that many alternative points for monitoring transaction flows are open to the national accounts compilers in estimating GDP and related national aggregates.

Thus it can be seen how

(1) The transactions summarised in production accounts of individual producers are shown as flows through the cost side, B and the receipts side, E;

(2) recording the aggregate of these flows through B or E gives two alternative ways of monitoring the total turnover (gross output) of the economy;

(3) total turnover (gross output) of goods and services during the period can also be recorded as the sum of flows of goods and services through A (as may be done in an interindustry study);

(4) if we deduct intermediate consumption from gross output we get GDP — analysed as at D, 'Gross Domestic Product at Market Prices'. When GDP is measured as the flows through D, the 'incomes' (cost) approach is being used.

(5) At C, we get the equivalent, 'Expenditure on GDP.' When measured as the aggregate of the flows through C, the 'expenditure approach' or final products method is being used.

(6) Measurement of GDP as the net flows of commodity inputs and outputs, represented in the diagram by E, is the 'production approach' or net-output method (sales and increases in stocks, less intermediate consumption). When production accounts for industries are consolidated to form the production for the domestic economy as a whole, intermediate transactions cancel out.

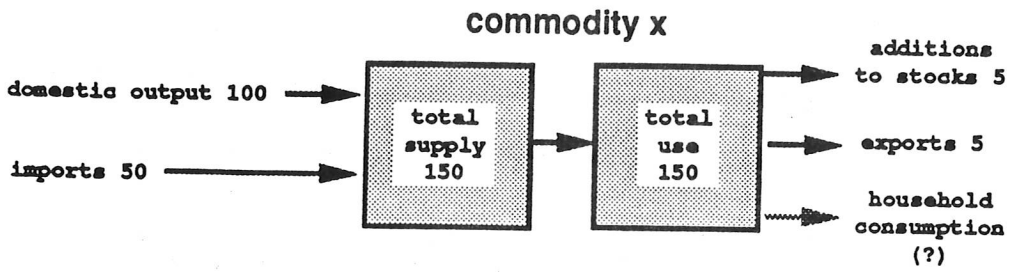
The relevant transactions for all of these approaches, will be recorded by both parties (transactors) to the transaction. Thus one can, in principle, collect the flows in GDP either from the producers (as costs) or from the owners of factors of production (as incomes). Similarly, one can ask the owners/consumers or the producers to report their purchases/sales.

For example, data on final consumption by households can be obtained from a household expenditure survey, a survey of retailer's sales, or be estimated indirectly by the commodity flow method — see diagram following.

In practice, national accountants use all possible channels to obtain the data. For example, some of the detailed components of household expenditure can be obtained reliably only from the householders, while others are best obtained from the retailers.

Some are best estimated by balancing the output of household-type commodities with the flows of imports, exports and changes in stocks using data obtained from producers and the Customs agency.

Thus, for a household consumption item X we may have data on the flows of commodities, as indicated in the following diagram, from which to deduce final household consumption of the commodity.



Given these figures of sources and uses of commodity X, its use in household consumption can be assumed to be the goods available from domestic output and imports, less the exports and additions to stock, i.e. 150 — 10 = 140 (note that allowances have to be made for trade and transport margins in practice)

This illustrates the *commodity flow* method. In countries with well developed economic statistics the national accountants will develop input-output studies and pursue to the limit a commodity flow method as a combination of the net output method and elements of the final products method. At the same time they may also employ the incomes distributed and final products methods at a broader level (eg without the commodity breakdown in the case of the latter).

Confrontation of data

The confrontation of the different estimates obtained by the various approaches and the effort to identify reasons for any discrepancies should result in progressive improvement in the quality of the estimates for the item concerned and, in a well integrated NSS, should lead to improvements in the basic statistics.

In practice, this process is not as rapid and comprehensive as might be hoped. Many developing countries may have no data at all on some of the major flows.

For example, household consumption is often obtained as a balancing item (which includes changes in stocks) from the balancing of GDP (estimated more or less crudely by the income or production approaches) against estimates of part only of expenditure on GDP (based on data available on government consumption, gross fixed capital formation and external trade).

Even countries which use a full range of statistical sources and all the possible compilation approaches are likely to be less than positive in reconciling the differences between alternative estimates.

reconciliation problems

Thus they may not be measuring exactly corresponding flows. The data collection timing may be different (some providing only delayed periodical benchmark data with which current indicators can be reconciled only a year or two after the first national accounts estimates are released).

Typically, components estimated via the incomes approach from income tax statistics will be some 2 years behind at least some of the data available for the expenditure approach and perhaps 1 year behind data from direct production surveys, which will provide partial data relevant to all three methods.

From the bottom up? or from the top down?

It may be noted that some of the methods of estimating GDP depend on building up from compilations based on statistical returns from individual economic agents, while other methods follow a *from the top down* approach, depending heavily on aggregated statistics from such sources as income tax, Customs data on imports and exports, and other government records. This latter approach may enable major components of GDP to be estimated directly, but without details of the contribution of individual economic agents or groups of agents being available.

Thus, the production approach, and also the incomes distributed method, when based on census/survey returns from producers, may give GDP estimates disaggregated in detail by industry and geographical areas.

When these are related to input-output analysis of commodity and interindustry flows there is the prospect of aggregate macroeconomic analyses of the production process being explicitly related to the basic statistics analysing activity at *meso* (intermediate) and micro levels.

On the other hand the income or expenditure approaches to compiling GDP do not yield such breakdowns.

For example, the external trade data used may give commodity detail supporting the input-output studies but the basic data do not normally identify the domestic parties to the import and export transactions.

Similarly, the commodity flow methods applied to estimating fixed capital formation and household consumption expenditure do not identify the specific transactors undertaking the expenditure nor do methods based on output or sales data from producing or distributing industries, such as the construction industry or retailers selling to householders.

Of course, when the various methods are used in combination, a considerable degree of disaggregation is possible and linkages with basic statistical systems can be established within constraints inherent in alternative sources and channels of basic information.

When such connections are made explicitly and positively, the problem solving capability of the total information system is increased very much beyond that of the individual information systems operating independently. Certainly the value of a system of national accounts becomes much more apparent when all the methods of estimation are being pursued from the bottom up as well as from the top down.

Input-output tables in the accounting process

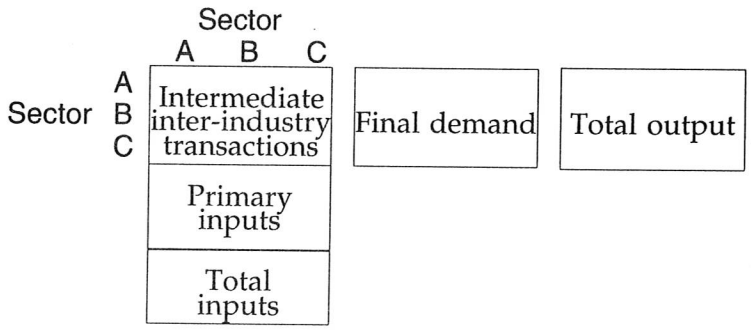
Chapter 12 introduced input-output tables in the context of the SNA's view of the *economy* as consisting of various categories of transactors which are alike in their economic functions and of *accounting* being essentially a cross-classification of transactions by type of activity and by nature of the transactor.

The input-output tables may also be an integral part of the process of compiling the estimates of Gross Domestic Product and of Expenditure on the Gross Domestic Product as they provide a valuable check on the annual national accounts, especially in regard to estimates of output classified by industry.

As the following consideration of a very rudimentary input-output table will show, the I-O matrix is, in essence, a disaggregation of the national accounts estimates designed to show the flow of production between industries, as well as the final outcomes of the process and its effects on the study of the distribution of income and its consumption or accumulation.

In fact, as we will see, the process requires the reconciliation of commodity flows, both for intermediate and final consumption, and makes the fullest use of the data sources involved in all three approaches to GDP estimation — production, income and expenditure approaches.

Input-output tables take the following basic form —



The quadrants are divided up into columns and rows for each commodity or industry sector.

inputs = outputs

It is axiomatic that the input into a commodity or industry sector must equal the output of that commodity or industry, ie., the column total for a particular commodity or industry group in an input-output table must equal the corresponding row total.

In the case of an industry input-output matrix, inputs include outputs purchased from other industry sectors, from imported supplies and also *primary inputs*, that is *compensation of employees, net operating surplus* and *consumption of fixed capital.*

In the following example there are three producing industry sectors, supplying intermediate output to each other, plus final output (goods and services for the final consumption of households, government and for exports; and capital formation by way of additions to stocks and additions to fixed capital).

Here we can see the destination of outputs by reading across the rows —

	Intermediate demand				Final demand				Total output
	Agr	Mfg	Other	Total	Cons. Exp.	Cap. Form	Exports	Total	
Sources of supply									
Agriculture	—	10	5	15	15	—	10	25	40
Manufacturing	3	—	17	20	20	8	2	30	50
Other	7	20	—	27	15	5	—	20	47
Sub Total	10	30	22	62	50	13	12	75	137
Imports of goods and services	2	5	3	10	3	2		5	
Primary inputs									
Wages and salaries	15	8	10	33					
Net operating surplus	12	4	8	24					
Depreciation	1	2	2	5					
Net indirect taxes	—	1	2	3					
Total inputs	**40**	**50**	**47**	**137**	**53**	**15**	**12**	**80**	

Thus $15 million of the output of agriculture goes to other industries, $10 million as purchases by manufacturing and $5 million by other industries. Following the expenditure approach for the economy as a whole, the expenditure on the gross domestic product is $65 million (ie the total of the final demand column $80 million after deducting the value of imported goods $15 million which represents the amount contributed by foreign rather than domestic production).

For this estimation it was not necessary to know the output of industry. We can arrive at the $65 million by collecting information directly on consumption expenditure, capital expenditure and exports less imports — the components of expenditure on the gross domestic product. Thus:

Expenditure on gross domestic product	*($m)*
Government Final Consumption Expenditure)	*53*
Private Final Consumption Expenditure)	
Capital Formation (increase in stocks + fixed capital expenditure)	*15*
Exports less Imports of goods and services	*-3*
Equals Gross domestic product	*65*

Alternatively, to follow the production approach, we do need the total output data and the intermediate purchases of each industry. Agriculture has a total output of $40 million and an input of intermediate goods and services (looking down the column showing purchases by agriculture) of $12 million. Its net output or value added is $28 million. The gross domestic product by this approach for all industries is the total domestic output $137 million less the total intermediate input of $72 million, giving the same $65 million that we get by the expenditure approach. Thus:

Gross domestic product by industry	*($m)*	
Agriculture	*28*	*(40-12)*
Manufacturing	*15*	*(50-35)*
Other	*22*	*(47-25)*
Gross domestic product	*65*	*(137-72)*

The balance on inputs and outputs is completed by the primary inputs which have to be added to the intermediate goods and services produced by domestic industry and to the imported goods and services. The primary inputs are the cost of the services of factors of production — compensation of employees for the factors labour and operating surplus (covering returns to the entrepreneurs and property owners). As we should expect, obtaining gross domestic product by the cost approach of adding the primary inputs gives the same $65 million as follows:

Cost structure of gross domestic product	*($m)*
Compensation of employees	*33*
plus Net operating surplus	*24*
plus Net indirect taxes	*3*
plus Depreciation	*5*
equals Gross domestic product	*65*

GDP at constant prices

We have noted the fact that extensive commercial accounting and economic accounting systems are possible because they can aggregate the many different kinds of transactions and the great diversity of goods and services involved using the common unit of market value.

Is the difference real?

These systems generally serve their purposes well enough in analysing activities within a given accounting period or changes from year to year or in making comparisons between similar economies. But they raise serious difficulties when the prices or qualities of goods and services transacted differ considerably over the time under consideration or between economies being compared. Even in the short term, a recorded increase in personal consumption by an individual or for the whole household sector may reflect nothing more than increases in the underlying prices (or declines in quality, equivalent to an increase in prices).

Of course, the transactions covered in GDP and related flows are accounted for in money values at the actual prices which applied at the place and time the transactions took place.

But the money value *unit* is itself subject to change as prices change. So, if users need to know the extent to which differences in economic aggregates are *real*, as distinct from being due to differences in the general level of prices, it is necessary to provide estimates conventionally described as being at *constant prices*.

The expression, *at constant prices*, is often used interchangeably with *quantum, volume, quantity*, and *in real terms.* It should be noted however, that measures at constant prices are not volume series; they are value series, adjusted to remove the effects of price changes. Thus we cannot have it both — ways current price *value* series serve us well in enabling us to aggregate diverse

goods and services. But while we might want to remove the effects of price change for purposes of analysis, we obviously do **not** want to return to pure physical quantities which cannot be aggregated because of their diversity.

is 'real' GDP really real?

So what we are doing in making constant price estimates, and generally in adjusting value series with price indexes, is simply providing artificial proxies for quantity aggregates that we cannot directly measure. Certainly 'real' expenditure flows are not 'real' in the sense of being made *tangible* by adjusting for price changes.

In practice, when some prices and quantities are likely to have risen and some to have fallen, what the economic analyst wants is to be able to say something about both the average change in prices and the average change in quantities. This can be indicated if there is sufficient information available about the price and quantity movements to be able to decompose the change in value expressed in current prices into a price component and a quantity component.

In principle, estimating GDP at constant prices involves expressing the current price value of each component transaction as the product of a price and a quantity and then substituting for each actual current price the corresponding price in the chosen base period.

Aggregates at constant prices for each period are obtained by summing constant price values for the transactions covered. The average prices of the base period, in effect, act as weights to combine the quantities of the individual commodities involved in the transactions. Thus, for a given base period, constant price measures, in theory, will vary only with changes in the quantities of the goods and services transacted.

'real' income?

This process of substituting base year prices can only be applied to the values of goods and services produced or consumed. It cannot be applied to incomes (eg pensions, interest and dividends clearly cannot be disaggregated into a price and a quantity).

It follows that constant price revaluations of GDP are applied to the transactions covered in either the net output approach (separately for inputs and outputs) and the final products method, for the various items of final expenditure. **The incomes-distributed method therefore applies only to current-price GDP estimation** and a full system of balanced economic accounts can only be drawn up in current price terms.

The two alternative measures of GDP at constant prices are thus —

(a) *industry-revalued* estimates, obtained by summing constant price estimates of gross product for all industries;

(b) *expenditure-revalued* estimates, summing constant price estimates of all final expenditures (plus exports, less imports).

Both are measures of commodity flows — recorded in the value adding process or in their final uses.

As this 'guided tour' concentrates on statistical frameworks and systems, it will not be exploring further the methodology involved in estimating GDP at constant prices. But it is worth noting that it involves quite complex procedures supported by extensive collections of data on quantities and/or prices of commodities and the construction of price and quantity indexes derived from these. For a description of the processes developed for this in a representative country see Australian National Accounts, Concepts Sources and Methods, Australian Bureau of Statistics, May 1990, Catalogue No. 5216.0.

Balancing items, residual discrepancies, accuracy and reliability

In the end, when the accounts (in current prices terms) are balanced up at national and sector level, judgements are likely to be made in adjusting component items down to the level of detail for which fully consistent accounts and tables can be published with reasonable confidence.

Accuracy, which relates the estimates to their true value, is virtually impossible to assess precisely for most national accounting aggregates.

We have seen how the double entry nature of the accounting system may be used to derive items, such as household consumption, as balances.

However, as far as possible the national accountant will try to produce independent estimates for each side of the accounts so that the balancing item is reduced to a measure of statistical discrepancy which will flag weaknesses in the estimates and will encourage development of sources and methods that will reduce the discrepancy.

Two points about the statistical discrepancy are worth noting

(1) the fact that it is shown by many countries as a balancing items on the expenditure side of an account should not be taken to imply that measurement errors are all on that side of the account. Its placement on one side or the other is usually a matter of convention. It may in fact be just as valid to record the discrepancy on the other side of the account or to average it. Some countries will acknowledge each of those views of where the residual errors may lie and will offer alternative figures of (eg) income based and expenditure based GDP.

(2) the size and systematic bias (positive or negative) of the discrepancy is no more than an indication of the quality of the estimates, since measurement errors in income components may offset errors in measuring the expenditure components.

For most uses of national accounts interest focuses on the movements in the figures, rather than on their absolute size. Fluctuations in the size of the statistical

discrepancy (or even a switch from a negative to a positive discrepancy, or vice versa) are therefore of much greater concern to users.

Beyond the national income and expenditure accounts

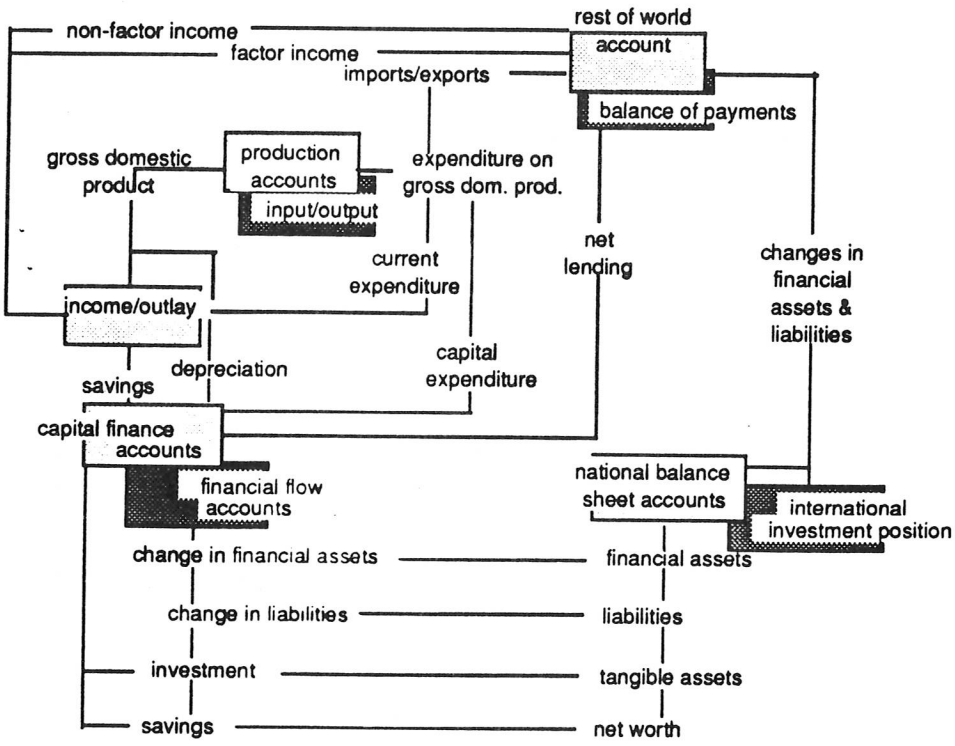
In illustrating the nature of economic accounting systems and their processes we have concentrated on the basic accounts of the system, as alternative measures of income, product and expenditure confront one another and have to be reconciled.

But this process of reconciliation of data sets goes much further than this, to involve the financial accounts and specialised statistical systems, such as those prescribed by the International Monetary Fund in their balance of payments and government finance statistics reporting systems.

A well developed integrated economic accounting system will then involve items in one such set of accounts being carried over to another in a systematic balancing up process.

Thus, as the following chart indicates, balancing items in one account, such as changes in financial assets and liabilities in the rest of the world's account of international transactions, will have to be reconciled with items in another (in this case, changes in the levels of financial assets and liabilities for this sector recorded in the national balance sheet account).

CONNECTIONS BETWEEN SUB-SYSTEMS OF A WELL-DEVELOPED ECONOMIC ACCOUNTING SYSTEM



Naturally there will be differences of opinion as to whether the figuring in one of the component systems should be adjusted to match that undertaken within another, or vice versa — eg is balance sheet data inherently more reliable than transactions data?.....and so on.

In the last resort the national statistician may have to make the judgements but the task is a complex and continuing one in which ground rules will be changing in the light of experience and improvements in sources and methods.

Concluding reflections on economic accounts and processes

These past four chapters have depicted the body of economic statistics as a system which is held together by three related characteristics

(1) The use of market exchange values as the common denominator for measuring economic activity.

(2) Consistency with well established business and government accounting conventions and records of activity in these money unit value terms.

(3) A basis in contemporary theory relating to economic behaviour and economic management in developed or developing market economies.

Both in the design of economic accounting systems, represented by the UN SNA and related systems, and in the practices adopted by economic accountants, there is the vision of a very large comprehensive data system and a dedication to a set of world standards adapted to individual country requirements in accord with their circumstances and needs. The discipline of the accounting process forces a high degree of coherence in statistical systems extending to those only loosely connected with the basic national accounts.

CHAPTER 15

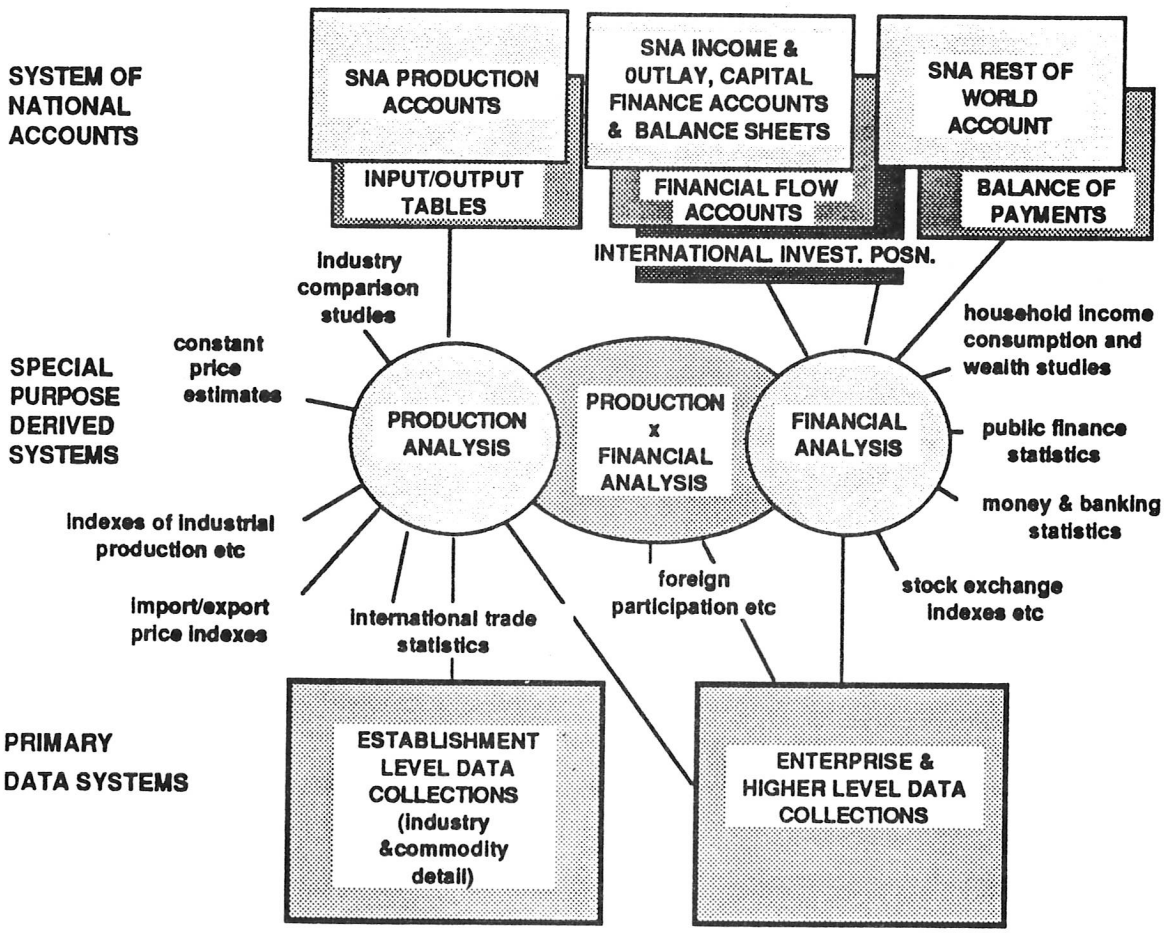
PRIMARY AND SPECIAL PURPOSE DERIVED ECONOMIC DATA SYSTEMS

This chapter shifts attention now to the characteristics of primary data systems which, while used to support the national accounts, may well have quite independent origins and a life of their own in supporting special purpose data systems.

Broad categories of economic statistics

The primary data systems, supporting the main special purpose derived systems and the national accounting system, can be represented in a simple two-way split, as follows —

PRIMARY & DERIVED ECONOMIC DATA SYSTEMS



Administratively the various statistical programs and projects may not be grouped in this manner. However it is a convenient framework for discussing systems which have common characteristics, in so far as they relate to the same levels of accounting.

General problems of system integrity

We have viewed the system of statistics available for the economic analyst as constituting a system to the extent that it is a more or less integral part of a total system of economic accounts as outlined in the SNA and its extensions.

what to do about inconsistencies?

We have noted however, that the national accounts' balancing process involves adjustments to the data from supporting statistical systems where different measures of approximately the same concept are inconsistent. Sometimes this leads to coverage and other adjustments to the data systems to make them fully consistent with the national accounts and with each other. More often this will not be possible because the statistical collection was developed basically for some other purpose or is based on administrative by product sources which are not under the national statistician's control or are subject to some inherent constraint.

1 — primary surveys
2 — adjustments to NA concepts
3 — adjustments in balancing up

In any case there tends to be a necessary distinction to be drawn, at least in operational terms (if not in principle) between the economic accounts system on the one hand and primary economic data systems on the other.

4 — feedback to primary data base?

Users can expect that a determined effort will have been made to make the former a fully reconciled, integrated set of statistics, generally conforming to the specifications of the SNA's definitional system.

They may not be able to assume this for the full range of primary statistics, even though they are important inputs to the national accounts. The national accountants may endeavour to make the necessary transformation explicit, but the user cannot always take such explanations at face value. This is because of the kind of **complex balancing up adjustments made at the aggregate level in successive rounds of preliminary and final publications for different parts of the total economic accounting system at different times (and subject from time to time to major long term revisions as new data sources or new methodology are introduced).**

Even where the primary data systems have been developed specifically to support the national accounts, the confrontation process may not result in survey results being modified as they are incorporated in the national accounts. **To the extent that the adjustments are substantially subjective judgements, or take place after publication of survey results, there will be a loss of comparability which may present difficulties for users of statistics who were not party to the adjustment process.**

NA — a 'test bed'

It is certainly important that everything be done to integrate survey systems and national accounting estimation processes as directly as possible. In particular, the confrontation of survey estimates with other national accounting evidence may throw up failures in the survey system which can be corrected in revised survey estimates — not infrequently a regular series *goes bad* and it is often the national accountants who will first recognise that something is wrong as they work with the series in association with other related data in their national accounting *test bed*.

There is clearly a general thrust in economic accounting to produce more and more detailed breakdowns of national aggregates, reaching down to the level of microanalysis available from the surveys of individual economic agents. This means a system-wide balancing up process which could extend to adjustments to the original survey results at the cell level and to modifications to the original questionnaires in order to fully reconcile the various components.

ultimate balancing-up objectives

With computer-based systems, it seems likely that it will in time be possible for the economic data held at all levels of aggregation to be used confidently as it becomes possible to reduce and correct for discrepancies and to render the derived and primary economic statistics more compatible.

It will therefore be useful to keep this general objective in mind, as we look at the characteristics of the primary data systems available to the economic analyst. We will accordingly group these systems in terms of the broad perspectives of the SNA and the two ways of dividing the economy into sectors depicted in Chapter 12.

Characteristics of access routes to source records

However desirable it may be to design the data access and processing system from the ground up, most statisticians face a situation in which there is already a considerable investment in existing collections and compilations.

inconsistencies in reporting requirements and processing standards

Too often the national accountant's problem is not with the state of the basic records and accounts that are the available data input but is with the inconsistent reporting requirements and processing standards of the various statistical agencies who collect and process the data for many separate purposes.

Thus problems of reconciliation in estimates made by the alternative expenditure, production and income approaches (evident in the size of the statistical

discrepancy items in the national accounts) can be traced to such problems as these:

- **the information gathered by the alternative approaches may not be for precisely the same economic entity,** eg where one set of information relates to the enterprise as a whole and another relates to only some of the establishments which comprise the enterprises. Even within a single continuous survey there may be uncertainty as to what statistical units of a complex enterprise are being covered in survey responses;

- **the information may not be from the same set of reconciled business records** — one enquiry may have been answered from management records and another from a (different) set of actual or notional records maintained for taxation purposes;

- **the separate enquiries may have been in respect of different accounting periods** — a financial year in the government agency administrative record and a calendar year in the direct enquiry;

- **the concepts, definitions, geographical coverage** etc applying in one enquiry may differ from those stipulated in another;

- **administrative byproduct data may only be accessible after aggregation** and in a classification system which is different from that being applied to data obtained directly from economic units;

- **reporting for administrative purposes may be biased and inconsistent with reporting for statistical purposes;**

- **differences in the timing and method of enquiry** (eg by interview or by mail) may elicit different reports in respect of the same record;

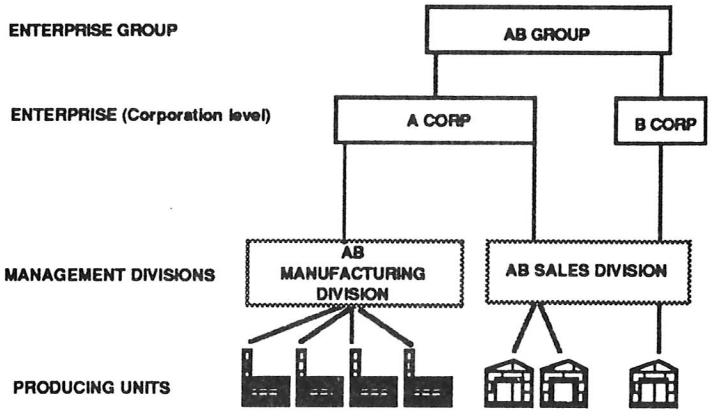
- **the report of one side of a transaction by reference to one party may differ from the report of the other side of the transaction by reference to the other party** because of differences in the ways in which these are brought to account (eg is it cash or accrual accounting? inclusive or exclusive of discounts, delivery charges etc?).

SNA's presumptions about statistical units & their records

Before examining the data collection implications of the SNA's sector accounting logic, it is worth considering first what is meant by levels of accounting unit. How would the accountant of a complex enterprise be likely to view this and how do we want the responding business accountant to report the business's activities?

A group of associated companies and their various divisions and physical places of business might have something like this four level accounting structure —

ACCOUNTING LEVELS IN A MULTI- INDUSTRY ENTERPRISE GROUP



Accounts would be kept for each corporation A and B (at least for income tax purposes) and for the AB Group as a whole (consolidated accounts eliminating transactions between companies of the group). There would be some records for each of the physical operating locations and probably more comprehensive accounts for divisions of the whole enterprise group.

The statistician is likely to be interested only in the producing unit (for which the standard SNA definition is given which will not in all cases coincide with the physical addresses depicted here). The statistician will also be interested in the two levels of enterprise type unit (which will be identified by standard rules which may not in all cases coincide with the perception which the group accountant has of the legal entities in his enterprise).

effect of classifying units by their predominant activity

Note that when the statistical units at each accounting level are classified to industry, they are classified on the basis of the predominant activity in each case. The effect of this is that we can get very different results from surveys, depending on the level of accounting for which data are reported.

Thus, in the case illustrated in the previous diagram, assuming that each producing unit has 100 employees, we will get the following different measures of manufacturing industry employment from surveys in

respect of a) establishments or b) corporations or c) enterprise groups

Manufacturing Industry Employment	
— in establishment statistics	400
— in corporation statistics	600
— in enterprise group statistics	700

It goes without saying that it is vital that such surveys clearly specify what statistical units are to be covered and that the respondent is in no doubt as to what the units are. A well maintained central register of statistical units is essential to provide positive control over a range of economic censuses and surveys.

Note that a business can be at once an establishment, a corporation and an enterprise group and would be recorded as each of these in a hierarchical register of business units. Note also that a household is a unit constituted by related persons in a dwelling and can be coextensive physically with a business establishment (as with a small shop).

The SNA presumes that it is both meaningful and practicable to record the flows of domestic production and expenditure between establishments, the flows of income and finance between enterprise type units and the flows of external transactions at the level of nations as a whole.

At the same time the logic of the system also requires that the flows recorded at these different accounting levels will all balance up when aggregated for the nation as a whole — eg that the production recorded by establishments is consistent with the income generated by them as recorded by the enterprise which owns the establishments.

Again, the savings item as it appears in the income and outlay accounts must be the same in aggregate as it appears in the capital finance accounts, even if the former be based on reports from enterprises as the statistical unit while the latter, perhaps, is based on reports in respect of enterprise groups.

business accounts — naturally consistent from level to level

There is no problem in the common case of the single-enterprise/single-establishment enterprise group (where establishment, enterprise and group are one and the same thing), because one set of records only is to be accessed, whether the standard statistical unit specified for the data collection is establishments, enterprises or enterprise groups.

For the multi-establishment corporation and the multi-enterprise group, normal accounting practice (in which components must add to totals and accounts must balance) would ensure that such records as are kept at each of the accounting levels of the enterprise group are

at least consistent and linked from level to level in the same way as in the SNA model.

To illustrate, the income and outlay accounts of a subsidiary multi-establishment corporation should record the elements of operating surplus generated by each of its physical producing units plus any receipts of non-operating income, such as interest and dividends due to the enterprise.

The records should also show how the enterprise disburses this income in various ways, such as by way of dividends or payments of interest to its parent company, or accumulates it as savings available for financing the purchase of real or financial assets.

These accounts may be brought together and further consolidated at the level of the enterprise group, at which level the parent company may undertake the necessary capital raising or investment for the members of the group.

The implication of the SNA design is that businesses do keep adequate branch accounts down to the level of their individual establishments.

Problems for a national economic statistics report system

are activities of component establishments fully accounted for?

Certainly, enterprise and enterprise group accounts are consistent in practice and the aggregate financial flows should be the same whether the enquiry by the statistician was addressed to the group accountant or to the accountants of each company in the group.

But, the statistician will not have the same confidence that the production data which he obtains in respect of establishments will be consistent with the income and outlay and capital finance reporting, particularly if the production survey enquiry was addressed to the managers of the producing units rather than to the accountant of the enterprise which owns them.

This is because the component establishments cannot be presumed to each maintain records which give the full accounting of their performance on the same basis as would be required for the enterprise as a whole.

The inherent weakness of establishment records as a source of statistics of industrial production is that transactions such as purchases and sales are contracts between enterprises, not between establishments.

Such transactions can be imputed to establishments, but since the profit maximising unit is the enterprise, there is no compelling reason for doing this for management or legal accountability purposes, particularly when it involves the allocation of common costs between several establishments and the imputation of purchases and sales between the establishments as if they were separate businesses.

However, given that we do want to be able to analyse the nation's economic activity in detailed geographical and industrial terms, we must try to identify transactions between physical production units, and we have to live with the data collection and estimation problems posed by the SNA. We must design basic data collection systems that will reflect its logic to the extent that it is practicable and worthwhile.

The question of what is practicable and worthwhile in a national system of economic reporting will depend on the circumstances of each country. In particular it will depend on the quality of the original records maintained by the business, governments and other institutions involved.

Those transactors who do maintain regular accounts can provide the statistician with an input of data which has already been summarised under the basic headings required by the national accountant. Thus, to the extent that the economic agents maintain annual accounts in accordance with the more progressive conventions of business accounting, national accounting would be substantially a simple process of aggregating by industry, sector etc the fully consistent and reconciled data presented in these annual trading and profit and loss accounts and balance sheets.

In practice, the statistical programmes tap the records of businesses at various points in their summarisation process.

At the one extreme, in overseas trade statistics, each individual import and export transaction is the input to the statistical system.

At the other extreme, company taxation statistics are based on the summaries of results in annual accounts. In between these extremes, are various censuses and surveys sometimes requiring the respondent to recompile the records of individual transactions to the statistician's specification.

work with accepted accounting conventions

As a general point, it is worth observing that the quality of the input of data to the statistical system is likely to be at its best if statisticians study the accounting practices of the businesses which are the subject of their censuses and surveys and tap the available business accounts at the most appropriate point and in terms that are in full sympathy with accounting conventions.

To illustrate, they should not simply ask for details of 'value of output' or 'value added' or some other item recommended in an international standards manual if the requirement can be met more easily by asking for the conventional accounting concepts (eg sales, purchases, and changes in stocks) from which the statistical concept can be derived .

The appropriate data collection timing, from the point of view of accuracy of reporting, is when the accounts for the year are complete and reconciled. But good reporting may be possible much earlier if the information required is of the kind that an accountant might be expected to present as estimates for current management purposes.

Ideally the statistician should impose reporting requirements on the economic agents which can be accepted by most as appropriate for their own accounting purpose as well as for the purposes of the industry to which they belong and for the national policy makers whose decisions form part of their working environment.

To the extent that the statistician works with the accounting fraternity and can influence their practice, the quality of the basic records is not entirely a 'given' condition for statistical programmes.

Real life units and standard statistical units

The problem of defining and implementing standard statistical units in data collection and processing operations is a particular instance of a problem which becomes more acute as information systems are required to serve a wider and wider range of uses and users. It is a problem which has always been a basic concern of statisticians — that of **resolving the great variety of real life arrangements into the more limited number of standardised perceptions for which numerical summarisations acceptable to information system users can be produced.** Essentially it is a problem of language.

business branch structures are individual arrangements

Accountants depict the accounts of their enterprise according to the way the organisations choose to be divided up for their particular management purposes. They are generally not concerned about how other enterprises are divided up for branch accounting purposes and so practices may differ widely even where physical arrangements are similar.

— but users seek meaningful summarisations at broader and broader levels

On the other hand, the statistician is trying to establish with accountants and economists a common perception of real economic activity, adequate for the working purpose of more and more users interested in broader and broader levels of aggregation, serving wider and wider purposes, to the point of requiring monolithic accounting systems.

To the extent that it specifies the main variables of interest to economists, as derived from items which have general currency in business and government accounting, the SNA and its supporting economic statistics approach this ideal of a unitary accounting system. **At least as a theoretical system it has established a common**

enough view of the full range of economic processes for data to be assembled and economic debate to be conducted using concepts and terminology which are now generally understood and accepted throughout the world.

However, notwithstanding the thrust of technology to foster interaction among greater numbers of people and to enable the integration of processes, we are still a long way from this ideal. **The fundamental difficulty remains of imposing the SNA's common set of perceptions of transactors and transactions and very wide range of purposes on accounting** *realities* **which reflect the non-standardised individual ways in which each enterprise organises its records in most countries.**

To summarise

This chapter shifted attention to the general characteristics of the data base of enterprise and establishment statistics and of the accounting records underlying these. It noted some formidable difficulties in the way of economists and economic statisticians as they endeavour to approach the ideal of a unitary economic accounting system, linking naturally with accounts maintained at the level of individual economic agents. We have recognised that basic differences in perception of accounting purposes are involved in this process.

CHAPTER 16

SYSTEMS BASED ON ENTERPRISE UNITS AND INSTITUTIONAL SECTORS

In this chapter we focus on the characteristics of enterprise and higher level data collections and the special purpose derived systems which they support, as indicated in the diagram 'Broad Categories of Economic Statistics' at the beginning of Chapter 15.

Institutional units and sectors

We are concerned here with the basic centres of decision in economic life — the owners of factors of production who control the production process, finance its operation, derive income from it and consume its final product. These are *institutional units* which can be categorised into *sectors* with distinctive economic functions.

These owning/consuming units are of the enterprise-type, such as corporations, corporate groups and unincorporated enterprises, and also governments, households and the non-profit institutions serving them.

On the basis of their principal function, such units are grouped to form the following institutional sectors (SNA Rev.4 titles) —

Non-financial corporations — *Institutional units which are principally engaged in the production of market goods and non-financial services*

Financial corporations — *Institutional units which are principally engaged in financial intermediation or in auxiliary financial services.*

General government — *Institutional units which, beside their political responsibilities and their role of economic regulation, produce principally non-market services for individual or collective consumption and redistribute national income and wealth.*

Households — *This sector covers all physical persons in the economy, grouped as households and including, in general, unincorporated enterprises owned by them. The principal function of households as consumers is consumption; as entrepreneurs, it is the production of market goods and non-financial services. The sector also includes non-profit institutions serving households, which are legal entities principally engaged in the production of non-market services for households whose voluntary contributions are the main resources.*

Each of the sectors is clearly different enough to warrant its own set of conceptual and operational standards because of the distinctive nature of its activities.

Thus, data systems which relate to households, for example, have a common focus on a homogeneous group of economic agents, so that these information systems are likely to complement one another (especially if common concepts and classifications are adopted) and are likely to be different from those relating to the economic agents in other institutional sectors which have basically different roles in the economy.

These broad categories provide a useful frame of reference for discussing different statistical collections and compilations as they logically fit into the national accounting system.

Certainly many of the primary data systems were established before the national accounts, and their development has followed an independent path as they have pursued their own specific objectives. But it is important to recognise potential linkages with other data for the same categories of statistical unit, which may benefit the specific applications as well as allow them to meet broader purposes.

In fact, many of the compilations of data relating to enterprise level transactions may not be readily aligned with SNA sectors and sub-sectors — some may only identify particular kinds of transaction, with the transactors inadequately identified.

International trade statistics are a major administrative by-product collection which illustrate this case. They relate to the transactions of domestic enterprise-type units, but computation systems generally only identify the flow of commodities into and out of the domestic territory (for each Customs port of entry by air, sea, road or rail). In this instance, breakdowns by sector, industry of enterprise or industry of establishment can only be *imputed*, very approximately.

enterprises can report financial and production data

Data collections relating to statistical units of the enterprise-type and aggregations of these will provide data for financial analysis and also for production analysis— value added data for broad industry categories made up of enterprise units and commodity data at national level.

Drafts for SNA Rev.4 are now emphasising the fundamental role of institutional units as the centres of decision for all aspects of economic life. Thus, to establish the links between production and the financial processes, they recommend a full sequence of institutional sector *current accounts* (production, distribution of income, use of income), *accumulation accounts* and *balance sheets*. Further analysis at the finer physical producing unit level of establishment-type units is treated as *complementary* and is limited to production and generation of income accounts for industries. As an establishment always belongs to an institutional unit it is possible to link the producing activities of industries and institutional sectors. Value added items will be cross-classified by sector and by industry.

Production statistics for enterprise level units are not well developed at present as countries have concentrated on establishment censuses and analysis by industry rather than sector. Enterprise collections have been mainly for finance statistics.

CLASSIFICATION OF INSTITUTIONAL SECTORS (DRAFT SNA REV. 4)

Non-financial corporations
- *Public non-financial corporations*
- *National private non-financial corporations*
- *Foreign controlled non-financial corporations*

Financial corporations
- *Central Bank*
- *Other depository institutions*
  - *Deposit money institutions* *
  - *Other* *
- *Other financial intermediaries, except insurance companies and pension funds* *
- *Financial auxiliaries* *
- *Insurance companies and pension funds* *

General government
- *Central Government (incl. social security funds)*
  - *Central government*
  - *Central government social security funds*
- *State Government (incl. social security funds)*
  - *State government*
  - *State government social security funds*
- *Local government (incl. social security funds)*
  - *Local government*
  - *Local government social security funds*

Households
- *Households, excluding non-profit institutions serving households*
  - *— headed by*
    - *Employers*
    - *Own account workers*
    - *Employees*
    - *Recipients of property & transfer income*
    - *Recipients of property income*
    - *Recipients of pensions*
    - *Recipients of other transfers*
- *Non-profit institutions serving households*

* separately for (1) public, (2) national private and (3) foreign controlled

As depicted above, the March 1990 draft of Revision 4 changes sector titles, drops *Non-Profit Institutions serving Households* to sub-sector status and extends the sub-sector categories. At the time of writing it represents the current international consensus on how countries might most usefully view their institutional structure.

Non-financial corporations sector — enterprise statistics

Non-financial corporations are enterprises which are mainly engaged in activities such as mining, manufacturing, retailing or the provision of services other than financial services. They may be corporate or quasi-corporate and be publicly or privately owned.

Note how the SNA Rev. 4 also suggests distinguishing enterprises subject to foreign control.

what really is the 'enterprise' when legal entities are associated?

The statistical unit in respect of which data are reported in many countries, is likely to be the family of corporate enterprises forming an enterprise group. While this approach to data collection avoids identifying formal transactions and links which are not economically meaningful, some countries prefer to use the individual legal entity as the basic unit for this sector and also for the *Financial Corporations Sector.* It is argued that this provides greater flexibility for classification by institutional sector, industry, and ownership characteristics.

is the legal entity fully operational?

In any case, the legal entity is not always an economic reality and may have to be combined with a legal entity which is operational in terms of undertaking a full range of transactions. See Bloem, *Units in National Accounts and the Basic System of Economic Statistics,* the Review of Income and Wealth, Sept.1990, for a discussion of this.

While interest may focus on this sector's role in production, income distribution and capital formation and financing of capital, *non-financial corporations* are also of interest in tracing the role of finance in the generation of incomes, savings and expenditure. Thus, to cover all sources and uses of money and credit, it is necessary to obtain direct or indirect information on sectors other than financial sectors because these other sectors supply funds as well as use them and not necessarily only through intermediary financial institutions.

financing as a secondary activity

Certainly many corporations which are primarily non-financial enterprises will lend surplus funds as a secondary activity to their main activities of manufacturing, wholesaling, retailing etc. In fact many retailers finance hire purchase and other forms of purchases on credit without recourse to a finance company.

Non-financial enterprises also, of course, maintain records of production, from which value added, capital formation, employment and commodity flow information can be provided for both the enterprise as a whole and for its component producing units.

We have noted that returns for the component establishments of multi-establishment enterprises are what is wanted for **industry level analysis of the production process.** But enterprise-level reporting may be all that is available in some countries.

In any case enterprise-unit industry data needs to be assembled, through one collection channel or another, **for the purpose of relating production to the financing of production** (since the latter is only capable

of being reported for owning/financing/consuming units of the enterprise type). **Ideally, enterprises would supply both their enterprise and component establishment level data, fully reconciled.**

company taxation returns

In addition to direct censuses and surveys of businesses and public enterprises, one of the major sources of data about non-financial enterprises is the income tax returns supplied by companies. It is particularly valuable since it is comprehensive in its scope — all companies report, regardless of their industry.

It is also comprehensive in its coverage of all the income earning activities of businesses. It is a potential source for data about income transfers but is unlikely to provide significant data on transactions in financial assets and liabilities.

The main problems to be faced in using this source are likely to be the lack of control over concepts and definitions of items on the taxation schedule, lack of a uniform reference period, processing time lags, and lack of access to individual returns. **Access for industry coding purposes and to positively align coverage with census and surveys is highly desirable. However Taxation Acts often preclude anything but aggregated data being provided.**

published accounts and registration Act reporting

Some analysis is often made of published accounts, generally using the family of legal entities (enterprise group) as the statistical unit. The main problem with this source is the limited coverage (covering only those enterprises listed on the stock exchange) and the lack of control over the information reported. Statements filed under *registration Acts* have potential and some countries do require enterprises to provide accounts in a standardised form — there are real possibilities for statistical analysis if the statistician can be involved in defining items, setting coverage and processing standards and defining statistical units to be covered — **desirably exercising some control via a central register of economic units.**

Financial corporations collections

Note that the terms Financial Institutions (SNA 1968) and Financial Corporations (SNA Rev.4) are being used interchangeably. The concept includes unincorporated enterprises primarily engaged in financial transactions in the market.

what is the role of financial institutions?

The nature of the role of the financial corporations sector and sub-sectors listed above is, generally, to promote production and other economic processes by **spreading risks and by facilitating movement of funds between savers and borrowers.**

information requirements for financial analysis

The data items that are of prime interest for financial analysis are those relating to the distribution and re-distribution of income, final consumption, savings, capital transfers, borrowing and lending, transactions in financial assets and liabilities and associated balance sheet and reconciliation data.

Enterprise level financial statistics usually include such programmes of basic statistics as the following:

- taxation assessment statistics,
- money and banking and private finance programmes,
- government finance statistics and
- household income and expenditure surveys.

Beyond these and their equivalents in different countries are many special purpose compilations — stock exchange activity, share price indexes and so on that are enterprise level statistics but may be somewhat peripheral to the measures of financial flows and stocks that are brought into focus by the SNA and related systems, such as are illustrated in the diagram *Beyond the National Income and Expenditure Accounts* in Chapter 14.

accounting legislation to monitor security

The nature of collections in this sector and of the special purpose statistical systems derived from them is likely to be determined by legislation, such as a *Banking Act* or an *Insurance Act*, designed to protect the public or to enable the government to exercise some control over the flow of funds within the economy and the rest of the world.

But data required for government regulatory and public protection purposes may not be what is needed for general financial and economic analysis.

The primary information requirements of the sector for regulatory purposes will relate to its activities of acquiring and disposing of financial assets and liabilities. Only secondary interest will be shown in statistics regarding transfers of income (except in the case of casualty insurance premiums).

In particular, regulatory Acts will not necessarily call for the classification of financial flows by the characteristics of the borrowers or lenders with which the financial institutions are dealing— the requirement is more likely to focus on the classification of their transactions and levels of financial assets and liabilities by type and by maturity

(eg long term and short term). Such data are needed to support their collective role in the operations of the domestic financial markets and in financial relations with the rest of the world.

Classification by the nature of the use to which loans are to be put (eg for housing) and according to the industry or sector, may subsequently be developed by negotiation with the institutions concerned.

At the same time the scope of reporting may be extended to cover other types of financial institutions (such as hire purchase and other finance companies, superannuation funds, building societies, short-term money market dealers, unit trusts and the like) which may be important and growing, but not yet encompassed by legislative controls.

Generally the development of statistics on the activities of financial institutions is likely to proceed by way of the involvement of the National Statistician in the design of the statutory forms required under the various controlling Acts or in arrangements to collect data to complement this information.

An integrated approach to this, by setting up some sort of industry working group, representative of all interested parties, may well be productive. An orderly development of financial statistics, which extracts the maximum value from the available information, is encouraged if the group sets the information requirements against the broad framework of national accounts— even if adequate flow of funds and national balance sheet statistics may be a long way off.

In the long run, the development of a flow of funds system is necessary to bring the myriad types of financial accounts and institutions together in a comprehensive and comprehensible statistical relationship. While this involves statistics from all sectors, the core of such a system will be statistics of financial institutions.

General Government collections

General Government includes agencies engaged in administration, defence, public order regulation, promotion of economic growth etc.

Government financial and trading enterprises are not included in this Sector. However the term *public finance statistics* may be likely to apply to the whole range of statistics about a *public sector* which embraces all of a government's activities in any of the SNA institutional sectors, except *Households*.

While the statistical unit for *General Government* compilations may be described as being of the enterprise type, the analogy with a private enterprise is not self-evident. The *enterprise-type unit* may be the whole central, provincial or local government — a huge entity with many different activities reassigned from time to time among various departments and other agencies. In some countries, as SNA Rev. 4 proposes, the social security fund may be sufficiently independent to warrant being separated as a sub-sector.

Some countries (eg Australia) define each department of government and each statutory authority as an *enterprise* (strictly an *enterprise-type* unit, analogous to an enterprise, as the term is generally used).

public finance statistics — a special purpose derived system

As a special purpose derived system, public finance statistics are likely to be concerned principally with financial accountability, summarising the accounts for budget review according to traditional classifications based on the particular structure through which each government is accountable.

Nevertheless many countries have also accepted the need to cast the data into an internationally comparable form. Pressure for this comes from the reporting requirements of the International Monetary Fund's *Government Finance Statistics* and from compilers and users of the national accounts who naturally want to relate the government's activities to that of the nation as a whole.

The *IMF Government Finance Statistics*' requirements have been sufficiently distinct from the SNA recommendations to be regarded as a special purpose system, rather than an integrated extension of the SNA. It is necessary, of course, to recognise that different forms of presentation of government finance statistics are necessary to enable it to be serviceable in the particular context in which they are used. This involves giving greater weight to the accounting practices specific to government administration than to the need to treat government transactions consistently with those of other economic agents.

It means adopting a standard budget classification, accepting the recording of receipts and expenditure on a cash basis and a general focus on the concept of budget deficit/surplus. The SNA and GFS data can be related conceptually and operationally and can co-exist as providing somewhat different but complementary views of the government sector.

consolidation of internal transactions

One of the complexities of working with public finance data is the need to consolidate transactions — eg when presenting an account of government operations for the various alternative sectors and sub-sectors, using data

drawn from the records of many different public agencies and funds.

Thus, for the purpose of rendering an account of the general government sector as a whole in relation to the other sectors of the economy, transfers made between funds and accounts within a government or between the central government and provincial governments are internal transfers and should be eliminated.

However, if the purpose is to show the operations of the central government in relation to provincial or local governments, such transfers should be added and shown as paid by one government and received by the other.

differences in accounting conventions

Similarly, the public finance statistics assembled for Budget and accountability purposes may not draw the distinctions between *financial transactions* and *real transactions* nor make the division of the latter between *current expenditure* (on goods and services) and *capital formation,* which are basic concepts in commercial accounting.

We have just used the term real, above, in the special sense in which the term is used in flow of funds literature. Like commodities and industries it has different meanings in different contexts. Flow of funds are divided into two parts, the real accounts and the financial accounts. The real accounts refer to the savings and investment decisions that flow from the production of goods and services that are the outcome of the income and outlay accounts. Transactions are real when they come from the side of the economy that produces something, as opposed to the financial side of the economy, which is the countervailing flow. In a different context, real is opposed to nominal and means something tangible, eg a growth in money values after being adjusted for price change (ie at constant prices).

Nevertheless this division of the real transactions into those on current account and those on capital account needs to be done for the purpose of relating government activity to activity in other sectors by way of a general economic accounting system.

what are the data sources for General Government?

The data sources for General Government sector statistics (as also for government components of other sectors) will generally be by way of government accounts — for example, budget papers, auditor-general's report, and treasury ledgers — published or made accessible to the government statistician.

For some purposes — eg in covering local government agencies and public enterprises — it may be expedient for the statistician to *short circuit* this process by direct data collections (perhaps incorporated in the censuses and surveys of private enterprises) or by reference to the central bank.

As already mentioned, perhaps the major problem in aligning statistics of the General Government sector with other sectors is that the transactions are **recorded on a payments basis** rather than on an accrual basis.

For example, the general government agency may not record the purchase of a good until a payment is made by the transfer of a deposit to the seller, even though there may be a period of, say, 6 months trade credit, evidenced by a bill of exchange or account receivable on the books of the seller.

This means that a sale by a private enterprise to a General Government agency may be recorded in one accounting period by the former and in another accounting period by the latter, giving rise to an element of the statistical discrepancy in the national accounts balance.

Private non-profit institution collections

We have noted that these institutions are expected to be downgraded to a sub-sector in SNA Rev.4, but they warrant discussion at this point because their economic role and behaviour in producing and consuming, and in raising finance and their statistical treatment are not very different from the General Government Sector just discussed.

Non-profit institutions serving household (NPIs) are not distinguished in the data collections of many countries because they are difficult to cover and are not thought to be significant enough to warrant special treatment.

However they are often important in developing countries, in particular, where religious missions are particularly significant producers of services.

In the case of missions and other religious institutions, it may be somewhat difficult to determine the enterprise-type unit unequivocably and to collect data accordingly. Record keeping will not generally be in accord with commercial accounting practice, given the non-profit nature of the operation.

no operating surplus element in NPI value added

As with General Government, the contribution of NPI's to GDP is assumed to be equivalent to the cost of producing their services to households, less the non-wage inputs used up in the process.

Since there is no net operating surplus element, the value added is simply the value of the labour input plus consumption of fixed capital and net indirect taxes, if any.

include voluntary labour?

In principle, this labour input should include an imputed value of voluntary labour contributed, as well as the wages actually paid.

In that case an imputed income matching this should be shown as being transferred back to the NPI — ie represented as supplementary to the actual money

contributions made by its members. In this way the full extent of the NPI's contribution to production and final consumption of services would be shown. However the SNA recognises the difficulties involved in measuring the labour inputs and imputing an appropriate wage rate.

Household sector

households — financial consumers and owners

The Household Sector includes all resident persons in their role as final consumers and owners of factors of production and also the unincorporated enterprises owned by them.

households as unincorporated enterprises

The latter's inclusion in this sector reflects the view that it is generally not possible to distinguish the household's financial activities and assets and liabilities in respect of its business from those in respect of the household (eg in the case of the farmer or corner shop where the owners of the enterprise live and work on their property). However such households are to be identified where possible in a sub-sector of the Household Sector.

Whether or not the household is also an unincorporated enterprise, the household is a unit of the enterprise-type in that it is an owner of factors of production (eg labour), a receiver and disposer of income and an accumulator of wealth.

The *household* is the *'person or group of persons occupying the whole or part of one housing unit and making common provision for food or other essentials of living.'* The SNA also recognises a more restricted definition, the *family household* where the multi-person household is limited to individuals related by blood, marriage or adoption who satisfy the other conditions of the housekeeping definition.

the individual

Also recognised in the SNA and its extension, *Provisional Guldelines on Statistics of the Distribution of Income, Consumption and Accumulation of Households* (Studies in Methods Series M No 61, UN New York 1977), is the individual person, as a desirable statistical unit for certain analytical purposes — eg in inquiries into factors underlying the distribution of employee compensation, entrepreneurial income or property income.

special purpose studies and national accounts

Studies such as these for most countries are special purpose ones, only loosely tied to the economic accounting system. A basic difficulty in relating the two is that the data sources for analysis of household income distribution, in terms of detailed categories of households, may be limited to direct surveys of the households themselves or taxation assessment data.

These may cover sources of income, expenditure and wealth that households are aware of and are able to report on. **But householders are generally not aware**

of all the elements of income and consumption that the SNA attributes to them — notably, withheld wage taxes and social security charges (which are part of compensation of employees as the employer perceives this, but are never money in the hand for the wage earner).

Other items of this nature are employers' contributions to private pensions, family allowance, health and other casualty insurance, life insurance and similar schemes, wages in kind and fringe benefits. **Such elements included in the SNA macro-economic analyses are likely to be excluded from distribution statistics because of the difficulty of estimation for different categories of households.**

Reconciliations should be possible, but are not easily effected, particularly in countries which are still producing very *broad brush* national accounting estimates.

Rest of the world

We have discussed the Domestic Sectors. In addition there is the Rest of the World which we discussed in Chapter 12. To complete our present exploration of Chapter 15's broad dichotomy of primary and derived statistics (relating to establishment level statistics on the one hand, and enterprise and higher level statistics on the other) we should note that the units for the Rest of the World are *countries.*

Transactions with the rest of the world are undertaken by all sectors of the economy, although the most important sectors are the *Financial Corporations Sector, Non-financial Corporations Sector* and *General Government Sector.*

The effect of a country's transactions with the rest of the world are initially summarised by the aggregate *net lending to the rest of the world,* while the overall viability of a country, after taking into account net acquisitions of financial assets and liabilities, is normally assessed in terms of movements in the official reserve assets of a country.

As the transactions entered in the rest of the world account are mirrored by transactions recorded in the accounts of domestic transactors, the sources of data applicable to the various domestic transactors are applicable here.

To sum up

The SNA, and its flow of funds and national balance sheet extensions, provide a broad institutional sector framework within which to pursue the development of enterprise statistics in an orderly fashion. This chapter indicates the nature and possibilities of the data available for each institutional sector. Statistics need to depict the sectors' different but complementary roles in the economy.

CHAPTER 17

ELEMENTS OF SYSTEMS BASED ON ESTABLISHMENT - TYPE PRODUCING UNITS

In this chapter we consider the characteristics of establishment level data collections and the special purpose derived systems which they support, as depicted in the diagram at the beginning of Chapter 15.

Establishment level production statistics

The general aim of production statistics is to record the key flows in the process of production — inputs and outputs in commodity detail, capital formation and employment and particularly, the net output or value added of each producer.

Thus the central objective is to monitor the contribution which producers (and in turn, industries and sectors) make to GDP as well as to enable the inter-industry flows underlying the process to be analysed via input-output studies.

Other data not directly required for economic accounting may also be collected. Thus data of a technical/ engineering-type (eg electricity used) or dealing with social/environmental impacts (eg emission of pollutants) are sometimes required.

A basic concern in integrating censuses and surveys of producers will be to ensure that the data are collected without duplication and across as wide a spectrum of industry as possible — a process which depends very much on operational arrangements centred on a central register of economic units and on the adoption of centralised procedures for the classification of establishments.

full, unduplicated coverage — homogeneous industries

We have noted that the preferred statistical unit for detailed analysis of production by industry and location is the establishment, which is less likely to have mixed activities than is the enterprise. We have also seen (Ch.12) that we need the establishment data, when compiled into industry aggregates, to be as homogeneous as possible.

Thus, the quality of production statistics is going to depend very much on the manner in which establishments have been delineated and classified to industries and this, in turn, involves both the design of the industry classification and procedures for coding each unit to industry class.

But, however well we perform the task of identifying and classifying our transactors, **they must be units for which basic production transactions can be reported.**

We will consider now the basic production transaction concepts which are common to production enquiries.

Basic production process concepts

These concepts relate to transactions that take place at the physical production level in the supply and disposal of goods and services — flows to and from such statistical units as farms, factories, shops, offices, and the results of these flows. The statistics should show **what** is being transacted, what kind of transaction (sale, transfer, etc), but generally will not indicate directly **with whom** (although it may be possible to infer this in broad terms from the nature of commodities transacted).

commodity and industry classifications are related

The importance of linking commodity classifications with activity (industry) classifications arises from the fact that industrial activity can be defined in terms of the products characteristic of that activity. This has been highlighted by the development of such publications as the UN *Classification of Commodities by Industrial Origin* (M.43).

The linkage of International Standard Industrial Classification (ISIC) with the Central Product Classification (CPC) and, through it, to the Harmonised System's (HS) tariff nomenclature, classifications of outlays by purpose (eg of Government) and other classifications relating to products, open up a wide range of possibilities for relating production to international trade, consumption etc. (see citations at end of Chapter).

data item concepts

UN recommendations on industrial census data items as developed for the purposes of the last decennial *World Programme of Industrial Statistics* in 1983 (see Statistical Papers Series M48 Rev.1 and Series M71) set objectives which fell somewhat short of what is needed for obtaining value added for the industries of mining, manufacturing, electricity, gas and water supply.

We will consider how census value added is derived from census concepts of output and input and how this relates to the true value added of these industries.

output can be at different stages of completion — valuation at producers' values or at basic values

Gross output. In general, this represents the value of all the goods and services produced during a given period, including work-in-progress and products for use on own account.

Valuation is usually at *producers' values,* that is, the market value at the establishment of the producer, or at *approximate basic values*, that is, producers' values less the relevant commodity taxes, net. (see SNA p.234).

Note that receipts for services done for others are only included if they are of an industrial nature (eg any receipts for rent or leasing of buildings or equipment would be omitted). The UN manuals assume that records of this kind are only kept at head office level in multi-establishment enterprises and that it is generally too difficult to cover by reference to establishments.

The SNA concept of *increase in stocks* poses difficult data collection problems — eg time of recording, valuation, consistency with *gross fixed capital expenditure* (reporting of work-in-progress on construction of fixed assets poses problems of valuation and ownership transfers) (see M.48, p 45 for a discussion on the time at which expenditure on fixed assets take place).

non-industrial service inputs may be missed

Intermediate consumption. In establishment collections, this relates to consumption of materials, fuels, supplies, etc and goods for resale and payments for industrial services such as contract and commission work and repair and maintenance.

net output or value added

Value added. This is the unduplicated contribution which the various industries make to the gross domestic product (GDP), defined as the *value of gross output less intermediate consumption.*

It will be apparent from the description of the elements of census gross output and intermediate consumption, that *census value added* is not true value added to the extent that any non-industrial services rendered by industrial establishments are omitted from their output and any non-industrial services used by industrial establishments are omitted from intermediate consumption.

The United Nations operational concept of value added in industrial statistics is, therefore, not net relative to the economy as a whole, but is only net relative to the industrial sector of the economy, the products of these sectors used up in production being the only inputs deducted.

Value of sales and receipts is the sales value of all goods for which ownership or effective right to use with a view to ultimate purchase has been transferred to others; and all services rendered during the period. *Value of purchases* is the delivered value of goods, intended for sale, the ownership or control of which has been acquired by the establishment. Value of *operating costs* is the cost of commodities consumed and services used during the reference period. *Gross margin* represents the difference between the value of the goods and services sold during the period and the gross cost of these goods and services (ie the value of purchases of goods for sale, adjusted for inventory change of the goods during the period. (See *International Recommendations on Statistics*

of the Distributive Trades and Services (E/ESA/STAT Series.M 57), pp 47-54).

linking—
(1) census value added with
(2) contribution to GDP (derived from the enterprise via the cost approach)

Components of value added. Value added has been discussed in connection with the production process and its derivation from data on production has been indicated. It is an important element also in studying the processes of distribution and consumption, since it is the source from which economic units derive the surplus to meet salaries and wages, depreciation, overhead expenses as well as profits.

Value added can be derived by working back from income tax data, starting with data on income reported by enterprises and working back to the contribution to gross product.

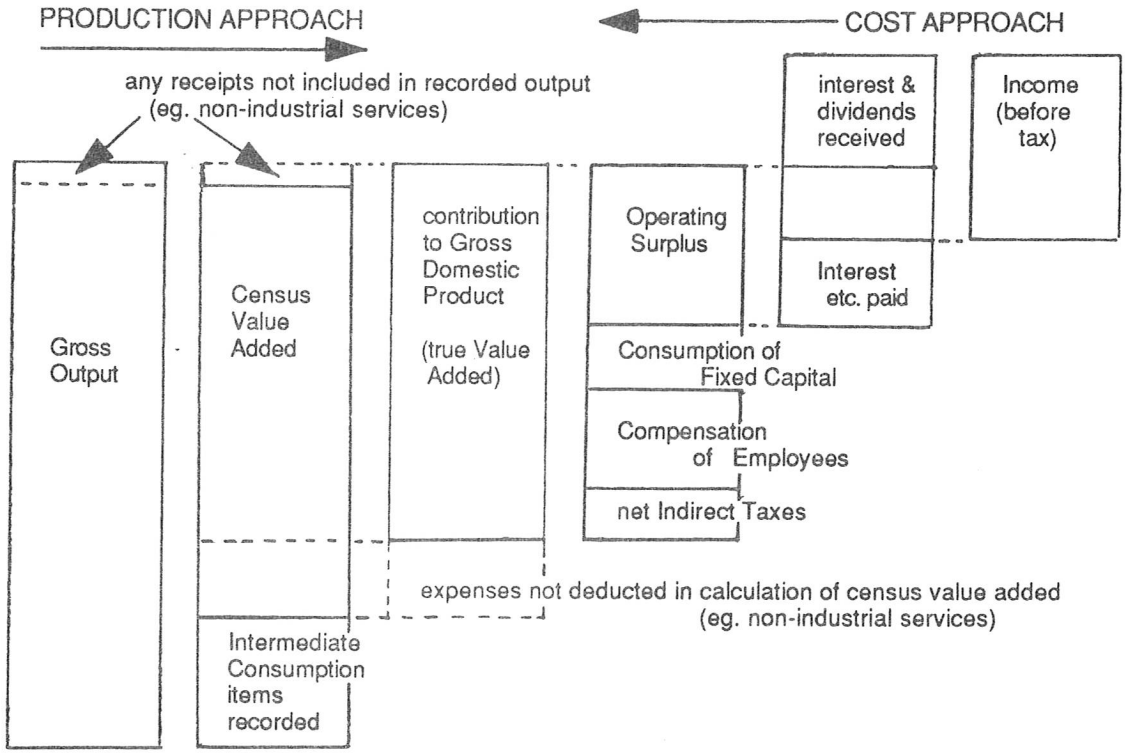
The net interest, rent and dividends are firstly combined with undistributed profits to derive operating surplus.

Then consumption of fixed assets, compensation of employees and indirect taxes less subsidies are added to operating surplus to derive the gross product to which the census value added measure approximates, the difference being non-industrial services which were included among the intermediate consumption items recorded.

Given the common origin in business accounts the two approaches should reconcile. However, they may be difficult to reconcile, mainly because of lack of standardization in the data collection definitions, differences in valuation, time of recording, record keeping practices etc. The definitions and reporting requirements are generally outside the statisticians' control, and they are not accorded an opportunity to help resolve inconsistencies until after the data have been collected.

The process can be illustrated as in the following diagram showing the conceptual links involved. It should be clear from the diagram that reconciling data from establishment censuses with data from enterprise accounts would call for some adjustment of the concept of *census value added.* In fact this may be only one of many problems in reconciling establishment and enterprise data at a national or sector level but it illustrates the need to look carefully at concepts (and operational practices as well).

CONCEPTUAL LINKS BETWEEN PRODUCTION APPROACH FROM ESTABLISHMENT DATA AND COST APPROACH FROM ENTERPRISE DATA



Note that the right-hand side of the diagram also indicates some of the income transfer process which occurs at the level of the business enterprise, as financial flows between institutional units occur outside the production process and determine the income available for distribution to the owners and to the government as taxation, after allowing for consumption of fixed capital and compensating employees.

Other key items of establishment data in establishment collections are *consumption of fixed capital, employment, compensation of employees* and *additions to fixed assets.* Their conceptual and operational definitions are detailed in the SNA manual and in the industrial statistics' recommendations of the UN Statistical Office.

Perhaps the most difficult theoretical concept to measure satisfactorily is *consumption of fixed capital.* In fact it is incapable of direct measurement and the conventional allowances made by businesses or allowed under income tax legislation are not consistent with national accounting principles (see SNA paras 7.22 to 7.24, pp122-123). Because of this uncertainty, balancing items in accounts including this item will be shown both *gross* and *net* in SNA Rev.4.

Registration and classification of establishments

Once identified by some form of census of economic units or registration process and checked for conformity with units standards, enterprises and their component establishments will need to be classified and coded accordingly in the register of economic units.

The great majority of establishments will be single establishment enterprises. Those which are units of multi-establishment enterprises will first be classified to industry and then any administrative offices or other ancillary units and also the owning enterprises will be assigned a *reflected* industry code.

Other classifications likely to be applied will be *size* (employment data are likely to be obtained in the registration process), *geographic location, type of ownership* (as in the case of the simple dichotomy between the public sector and the private sector, or the more elaborate SNA institutional sectors) and various other characteristics needed for analysis or operational control.

Such classifications can, of course, yield cross classifications for the purpose, for example, of providing data by industry by type of ownership by size of establishment.

Nature, purposes and principles of industry classification

Industrial classifications, such as ISIC and country classifications based on this are **designed primarily as a system for classifying establishments,** eg individual mines, factories, shops etc., by industry.

They may be used for classifying other statistical units, such as enterprises, but the concept of an industry as a kind of activity becomes less meaningful the broader and more mixed the activities of the unit being classified. Consequently statistics of enterprises classified to industry are usually published only for the broader levels of an industry classification — it is suggested that only ISIC Tabulation Categories and Divisions will be useful for institutional units.

just what is an industry?

Thus an industry — ie an industry class, group etc in an industry classification, **consists of the units which have been classified to it.** This fact is often lost sight of in debate.

An industry classification is sometimes confused with a commodity classification in which similar commodities are grouped according to their physical characteristics, use etc.

It also needs to be distinguished from a classification of occupations in which similar occupations are classified according to characteristics of the work they entail.

Key points to grasp are—

- **an industry is not all the units engaged in producing a particular commodity and**

- **the units forming an industry do not only produce the class of commodity by which the industry is described.**

Thus it may be said by some that all factories engaged in manufacturing agricultural tractors are in the agricultural machinery and equipment manufacturing industry. At the same time, if a factory engaged in manufacturing agricultural tractors is also engaged say, in manufacturing bulldozers, it might be said, in this sense of the term, that it was also in the construction and earthmoving machinery and equipment manufacturing industry.

However, in an industrial classification, each unit has to be classified uniquely to one class, so that only factories whose major activity is manufacture of agricultural machinery and equipment can be brought together to comprise an agricultural machinery and equipment manufacturing class and only those whose major activity is manufacture of construction and earthmoving equipment can be brought together to comprise a construction and earthmoving machinery and equipment manufacturing class.

Each *industry* is defined in terms of a specified range of activities, designated as primary to it. Similarly, each group is defined in terms of the activities designated as primary to the classes within that group, and so on. An establishment which is mainly engaged in activities which have been designated as primary to a particular class is classified to that class, whether or not the establishment is also engaged in other, ie secondary, activities.

It might be expected that an industrial classification, such as ISIC, would be identified in its listing of commodities primary to it, by groupings of commodities equivalent to those which might appear in a commodity classification, such as the Customs Cooperation Council's nomenclature for classifying goods in Customs tariffs. But in identifying an industry classification, the relevant question is **what commodities are typically produced by establishments? <u>not</u> what commodities are similar in appearance or use?**

As an illustration of this distinction, abrasive coated papers and plastic film sheeting, considered as commodities, might be regarded as having few attributes in common, and it might be considered inappropriate to bring them together in one category in a commodity classification. However, if the activity of producing abrasive coated papers and the activity of producing plastic film sheeting were commonly carried on together at the one establishment, the two activities would be designated as primary to the one industry in an industrial classification, despite the dissimilarity of the articles produced by the respective activities.

industrial origin commodity classifications show commodities under industries to which they are primary

One type of commodity classification in which dissimilar commodities may be grouped together are so called *industrial origin commodity classifications.* In an industrial origin commodity classification, commodities are grouped according to the industries in which they typically originate, ie the industries to which the activities of producing the commodities are designated as primary.

Thus the broad structure of an industrial origin commodity classification consists of industry of origin headings, and detailed commodity items are shown under each industry of origin heading to cover the types of commodities which typically originate in that industry.

Leaving aside the problem of treating commodities which are subject to overlaps between industries, each individual commodity item in an industrial origin commodity classification is usually shown only under one industry heading.

Of course, a proportion of some commodities is produced by industries to which their production is not primary. This is shown in tabulations of commodity data by industry of production, ie industry of the producing unit. However, such tabulations should not be confused with tabulations of the same commodity data by industry of origin (in accordance with a particular industrial origin commodity classification). These show the commodity data according to the industries to which their production is primary (not according to the industries in which their production actually took place).

industry classification should be used widely in statistics and debate

The main purpose of a kind of activity classification, such as ISIC, is to provide a standard framework for classifying establishments and other statistical units by industry in official statistics.

If it has been developed as part of an integrated statistical system, it provides for each individual establishment (or other statistical unit) to be classified to the same industry in all statistical compilations in which it is included. Desirably they should be used in economic censuses and surveys, population censuses and surveys, and in other statistics (national accounts, etc), derived from the basic statistics.

The assessment and application of an industry class

While ISIC serves as a world standard, it is desirable that countries prepare their own versions which will be consistent with their industrial structure as well as convert to ISIC at some level of its classification.

homogeneity of output (specialisation and coverage ratios)

Thus, in devising industry classes, the aim is to have classes relate to groups of establishments mainly engaged in the same or similar kinds of activity and which represent realistic and recognisable segments of

industry, ie industry classes should meet quantitative standards relating to homogeneity of output (in terms of minimum acceptable specialisation and coverage ratios) and importance (eg in terms of size and user interest).

industry coding principles

There are three basic principles for classifying establishments to the cells of an industry classification:

- At each level (eg division, group or class) an establishment can be classified to only one cell (eg a particular division).

- The division, group and/or class to which an establishment is classified must be related by aggregation or disaggregation.

- Each establishment is to be classified to cells according to its major activity.

data for coding

In determining the industry classification of an establishment there are basically two alternatives

(1) The respondent's own description of the activities of the establishment and their evaluation of the relative importance of the activities of the establishment

(2) Quantitative information (eg value data) relating to the kinds of goods produced or handled or the kinds of services provided, from which the relative importance of individual activities can be deduced.

In the case of the respondents' own assessment, whilst they will undoubtedly have a more intimate knowledge of their business than the statistician could attain, it is inevitable that different respondents will have different criteria in mind in assessing importance, and the weight which they give to each activity will differ from respondent to respondent. In classifying on the basis of quantitative data it is possible to avoid such inconsistencies in approach.

in principle value added is the best measure of level of activity — but is it available at commodity level?

The choice between various types of quantitative data depends largely on the practical question of what information is available for individual activities within all (or most) establishments in a particular collection.

Commonly this consideration has led to the adoption of value of gross receipts or output as the measure for assessing the major activity of establishments in integrated economic censuses. However, value added has also been adopted where it is available. In some industries other measures such as wages and salaries,

employment or respondent's description may need to be used.

In determining the industry classification of an establishment there are basically two alternative methods:

step by step industry coding

- Classifying it to a cell at the broadest level of the classification in the first instance and subsequently to cells at successively lower levels (eg classification to a division, then to a group within the division, and so on until the establishment is finally classified to a class or sub-class). For convenience this method is referred to as the *step-by-step* method.

direct coding

- Classification directly to a cell at the lowest level of the classification (eg direct to a class of the ISIC).

The step-by-step method of classification is generally adopted where quantitative measures are used to assess major activity.

reflected coding for ancillary units

Administrative offices and ancillary units are assigned a *reflected* industry code, generally corresponding to the industry which represents the predominant industry of the establishments administered or served by the ancillary unit.

coding of enterprises

Enterprises and enterprise groups are assigned a *reflected* industry code, corresponding to the industry which represents the predominant industry of the establishments owned and operated by the enterprise or enterprise group. The method used in classifying enterprises and enterprise groups should involve the following:

- Weighting each establishment (in the enterprise or enterprise group as the case may be) by the establishment's total value added or some substitute weight, such as total employment.

- Determining the predominant industry of the enterprise group and of each of its enterprises separately on the basis of the predominant industry of the establishments owned and operated by them in each case. This predominant industry of establishments is determined and ascribed to the enterprise or enterprise group at successively finer levels of the industry classification.

The geographical dimension and regional accounts

establishments are location-specific units

While the topic of regional accounts has not been covered in the earlier discussion of national accounts, at this point it might be noted that regional accounts and statistics would depend heavily on establishment data. The establishment is also the appropriate unit for geographical data on production, employment and sales data of special interest for market research.

cost approach with enterprise data raises problems

Generally, the cost (income) approach to GDP arising in each region of a country founders on the fact that such gross operating surplus data as is available (eg from income taxation sources) is in respect of enterprise-type units, many of which have activities extending across regional boundaries. For some industries, the best that can be done is some fairly crude allocation of national totals.

expenditure approach — how to cover imports and exports?

The expenditure approach is not used because of the absence of any regional counterpart of the statistics of external trade in goods and services obtainable from Customs administration.

To the extent that establishment data relates to physical locations which can be classified by State, there should at least be the basis for a regional breakdown of GDP and of income arising from this.

In principle, what is involved is to define resident transactors by adapting the SNA conventions to the regional situation. In fact the problems of classifying mobile facilities to regions (eg ships, aircraft, railways, pipelines, fishing boats, oil and gas rigs and the like) may not be analogous to the extent that aspects of national sovereignty may have no regional counterpart.

The problem will be aggravated in the case of a number of industries, for which, realistically, it is necessary to define the establishment unit as the enterprise/industry unit — all the operations of an enterprise for that kind of activity conducted at or from all of its locations in the country.

As will be discussed in Chapter 18, focussing on the characteristics of statistics of different industry groups, this is likely to affect particularly such industries as transport and communication, defence, and research and scientific institutions.

Statistical programme groupings

statistics organised on industry policy lines — potential for duplication when establishments and industries have secondary products

Data collection programmes tend to be organised in parallel with the way in which the economic agents in the economy are organised as producers and consumers of commodities. Thus a statistical programme collecting information about farm units may be conducted by a Department of Agriculture while another, collecting information about mining establishments, may be conducted by a Department of Mines and so on.

The arrangements for collecting production statistics thus tend naturally to be in line with very broad categories of an industrial classification.

This is true even when the collection of statistics is largely centralised in a central statistical office. In this case, however, it is likely that the programmes have a wider scope — eg, there may be a division dealing with industrial statistics, with separate sections of this dealing with the specialised aspects of mining statistics as distinct from manufacturing statistics.

There are advantages in this second more comprehensive approach, particularly when, with the development of a more complex economy, the multi-industry business becomes more significant. In that case problems of avoiding duplication and gaps in reporting the activity of their component establishments become harder to deal with.

The two broadly different ways of tackling this kind of problem at the operational level are —

Separate approaches

- most Western countries have developed establishment censuses and surveys separately for different industries;
- some have relied on enterprise data from taxation agencies and from direct surveys of enterprises;
- some have used both approaches and have tried to reconcile the two at the aggregate level;
- many gaps, duplications and inconsistencies in concept and timing have prevented any positive reconciliation.

Integrated approaches

- some countries are undertaking economic censuses in which enterprises supply data for the enterprise as a whole, plus detailed data for each of their component establishments and ancillary units with the two sets of data reconciled;

- this ensures complete coverage of establishments and ancillary units like head offices;

- it enables adjustments to be made to give true value-added at establishment or kind of activity level;

- it requires a good central register of enterprises and their establishments and a common processing system for all production censuses that are to be integrated.

industrial censuses need to cope with vertical integration of eg mining and manufacturing

Generally this kind of problem is most severe across those industries in which there tends to be a vertical integration of activity. Thus there will be problems in conducting separate censuses where (say) it is common for a manufacturing company to own its own mining establishments and wholesaling establishments.

Of course it may be possible to collect both establishment and enterprise data simultaneously from enterprise head offices in order to achieve complete unduplicated coverage of production and to adjust the census value added data to give true value added at establishment level.

However, to do this regularly in industrialised countries, across the full scope of the *International Standard Industrial Classification* (ISIC), raises massive logistical problems.

Even Australia, which has a centralised statistical service and has been dedicated to the concept of integrated censuses and surveys since the late 1960's, stopped short of this.

Its rolling annual establishment and enterprise level census program has covered all enterprises with establishments in agriculture, mining, manufacturing, electricity, gas and water supply. Other industry sectors were added to the scope of the integrated censuses in different years — eg construction in one year, distribution in another, so that such industries could be fully covered by census or sample survey only about every 5 years and personal and business services at longer intervals.

More recently, the ABS has introduced an economy-wide short-form survey, overlaying a less-intensive rotating census programme which obtains the commodity detail.

International trade statistics and commodity data

We have noted that we look to establishment level data collections for a detailed coverage of the inputs and outputs of commodities. The other major source of such data is international trade statistics — imports and exports of data by country of origin and destination in

relation to the domestic economy as a whole (ie without identifying the domestic transactors involved).

We deal with international trade commodity statistics here because of the close association they have with industry statistics and input/output analysis of commodity flows. But of course aggregate imports and exports is also a basic source for the financial flow and balance of payments analysis discussed in Chapter 12.

The source of international trade statistics is the administrative records of Customs agencies.

The statistical data item to be recorded is the individual import or export entry, so that statistical compilation is massive.

A very great deal will depend on the degree of cooperation between the customs administration and the statistical agency — particularly as to the extent to which the customs legislation and procedures have regard for the standards of definition and classification that are all important for the purpose of ensuring international comparability and of meeting the requirements of economic analysis.

The standard concepts and definitions which apply to the statistics to be derived from these administrative data sources recognise the limitations of the Customs documentation and coverage which, of necessity, stop short of meeting SNA and Balance of Payments requirements.

In addition, the scope of goods to be covered as imports and exports in the economic accounts, is defined to take in some items (such as transfers of ships and aircraft which occur outside the national boundary), which are excluded from the definition of *merchandise* to be covered in the regular international trade statistics. On the other hand, some merchandise movements have to be excluded because no change of ownership takes place (eg equipment on lease).

It is in the nature of international trade statistics that what is recorded as import or export transactions are the movements of goods through a Customs station or across a national frontier.

On the other hand, the national accounts and balance of payments systems aim for consistency with business accounting, in which the transaction relates to the change of ownership of goods imported or exported.

As this change of ownership may not necessarily occur in the same accounting period as when the goods pass through Customs (ie when finalised documentation enters the statistical system), trade data should, in principle, be adjusted to the change of ownership timing. This is

difficult to do, apart from the major identifiable items such as imports of aircraft and primary product marketing board consignments for sale overseas.

In addition to the problems indicated above, there are many others such as those relating to valuation practices for exports, which may be inconsistent with the valuation practices used for the imports of the country of destination.

Partner country comparisons may be further complicated by inconsistencies in the definitions and classifications of commodities and of country of origin and destination. Nevertheless, international trade statistics are at least the subject of a great deal of attention to ensure international comparability and there is close cooperation between the Customs Cooperation Council in Brussels, the International Monetary Fund and the UN Statistical Office to develop and implement common standards and classifications worldwide (extending to a direct linking of tariff, trade and production statistics classifications).

To summarise

We have looked at some of the key transaction concepts which will be relevant across all industries and explored the principles of industry classification that underlie establishment statistics for industry and regional analysis. We noted that the popular concept of industry is not the concept in industry classification. The nature of this concept needs to be understood by anyone working with industry statistics. (See especially International Standard Industrial Classification of all Economic Activities, UN, New York 1990, Statistical Papers Series. M 4, Rev 3).

We noted that establishment statistics are the main source of commodity data along with international trade commodity statistics which are based on Customs administration documentation. Of general interest in this connection are:

(a) International Trade Statistics, Concepts and Definitions, Statistical Papers Series M 52, Rev 1.United Nations, New York, 1982.

(b) International Convention on Harmonised Commodity Description and Coding System HCD or HS) Customs Cooperation Council, Brussels 1983

(c) Final draft of the Central Product Classification, (CPC) United Nations, New York, Provisional Statistical Papers Series M 77, 31 August 1988

(d) Standard International Trade Classification, Revision 3. UN, New York, 1986 Statistical Papers Series M 34, Rev. 3,

CHAPTER 18

MAIN FIELDS OF ESTABLISHMENT LEVEL STATISTICS

In this final chapter on economic statistics we discuss the special characteristics of the main fields of statistics in which production data are collected at the establishment level.

The way in which international manuals on production statistics are currently organised suggests the common groupings of statistical programmes in the field of production statistics. These are:

- *agricultural statistics (including forestry, fishing and hunting)*
- *industrial statistics (mining, manufacturing, electricity, gas and water supply)*
- *distribution statistics (wholesale and retail trade and associated services)*
- *construction statistics*
- *transport and other service statistics*
- *government services (public expenditure statistics)*

Most of their common elements have been discussed in the previous chapter. In this chapter we consider what is distinctive about each of these fields.

Agricultural statistics

The largest part of agriculture is in the hands of small unincorporated household enterprises (and small private companies in some countries) and the problem of the multi-industry enterprise has not generally been a dominant concern.

Significant household industrial-type activities may be taking place in the farm household, but these can be treated as secondary activities of establishments which are primarily agricultural product producers.

how to cover and value subsistence production?

In developing countries, the main conceptual problem in agricultural production statistics is how to measure the substantial volume of production for subsistence, rather than for sale.

There can be severe problems in covering own account production/ consumption of foodstuffs and other commodities and, in some areas, an even more intractible problem of imputing a price for such unmarketed products when an appropriate market price is difficult to find.

when is production occuring?

A further characteristic of agricultural and livestock production is its seasonality — agricultural statistics are usually recorded on the basis of the agricultural year which, especially in the southern hemisphere, does not coincide with the calendar year. Adjustments then have to be made to allow for the fact that major costs are incurred in one calendar year in respect of a harvest which takes place in the following calendar year.

the statistical unit

The customary establishment type unit for agricultural and livestock production is the farm holding.

The UN defines the holding as: "An economic unit of production under a single management comprising all livestock kept and all land used wholly or partly for agricultural purposes. No account is taken of the form of tenure or legal title to the land or livestock. A holding may be managed by one or more people or households or by institutions such as cooperatives, legal companies or the government. It may be made up of land located in a number of separate parcels, provided that all the land is operated as a single unit using the same labour, machinery and other inputs."

Notwithstanding the extensive nature of agricultural and livestock operations, the holding reasonably matches the SNA's ideal concept of an autonomous entity which engages in one, or mainly one, kind of activity at a single location (although *location* has to refer to the large areas for which agricultural data are normally published).

indirect estimating

Agricultural statistics usually involve mixed indirect estimating approaches.

Firstly, there may be the agricultural census or survey which (in deference to the generally low quality of farm financial accounts) normally collects only physical quantity data on output, which may be supplemented by information from marketing agencies.

Using price data and special surveys of farm costs, estimates are then made of contribution to national product.

Whereas, with other sectors, the need for price data may be largely to enable the adjustment of value flows to a constant prices basis, prices information for this sector is sought particularly to convert available stock and flows of quantities of inputs and outputs into values which can be aggregated and related through accounting identities.

Characteristically, much of the available data are associated with regulatory and economic support functions of government.

In many countries, regular data collection is confined to the large plantations which grow crops for export. It has proved much more difficult to cover the rural small holders, working land under traditional village land tenure to produce export crops for sale and other food crops for subsistence and sale.

The international standards for agricultural statistics are generally laid down in the programme manual for the decennial *World Census of Agriculture*: published by the UN Food and Agriculture Organisation (FAO). This aims for consistency with the SNA.

basic record-keeping deficiencies

The objective of covering all fields of production in a consistent fashion in terms of the unduplicated output of establishment type units should lead to the collection of data from agricultural establishments on a basis similar to that of other sectors of industry. But agriculture in most countries is particularly difficult to cover in this way because of the relatively poor level of farm accounting.

Historically agricultural statistics have tended to stand on their own, but with increasing attention being given to the requirements of economic as distinct from agronomic analysis.

Industrial statistics — mining, manufacturing, electricity, gas and water supply

The boundaries of the field have been defined by the UN Statistical Office for the purpose of such Statistical Papers as *International Recommendations for Industrial Statistics* (Series M48) and various papers associated with the decennial *World Programme of Industrial Statistics* (eg M71) and with indexes of industrial production.

boundary issues

One common element is that these are the industries which are essentially suppliers of goods, rather than services, although the distribution (as distinct from production) of electricity, gas and water is a service function. They need well integrated data collections and conceptual standards because their separate activities commonly occur at the same production site.

Another reason for the grouping is that large enterprises tend to spread across these fields, with the strong possibility of gaps and duplication in coverage and inconsistency in reporting (eg of transfer values where mining products are transferred to manufacturing enterprises of the same enterprise). Thus **the case for integrated enterprise/ establishment collections is very compelling.**

industry classification problems

The logical distinction drawn between mining and manufacturing is essentially a functional one, attempting to separate (a) the process of extracting primary products from their natural environment, from (b) the process of transforming materials or components into *secondary* products.

Electricity and gas production are also processes of transformation, although water supply could well be thought of as extracting rather than transforming. The common attribute of electricity, gas and water supply is that it is difficult to separate their distribution from their supply.

This distinction between extracting and transforming has some theoretical appeal but the distinction is less satisfactory in a situation in which some mining and manufacturing operations are highly integrated processes (eg the added value arising from domestic smelting and refining of metallic ores and concentrates). The ISIC treats this as secondary activity, but a national classification reflecting the reality of the country's industrial structure may well include this in the concept of a *mineral industry*.

Even when transporting intervenes between the extracting and transforming processes, development planners and others may legitimately identify (say) an integrated *iron and steel industry* embracing both mining and basic iron and steel products and require coherent statistical data.

The point for statisticians to consider is that, while standard systems of classification and aggregation are an important element in producing statistical materials under uniform and controlled conditions, they must be used in a flexible way to provide for alternative information requirements.

Data item concepts for mining, manufacturing, electricity, gas and water supply industry statistics can and should be treated in a uniform way — at least with regard to the basic accounting data relating to input, output, value added, capital formation, changes in stocks, and employment, wages and salaries and the like.

Naturally the classes of commodity appropriate in each case will differ and the standardised data collected in comprehensive censuses and surveys will need to be supplemented through special studies and enquiries. These might provide the engineering type data appropriate to each for the purpose of planning particular industrial projects.

In the case of mining, the quantities and values of individual minerals produced are recorded on the basis of the form in which the minerals are despatched from the working. For example, a metallic mineral is recorded as an ore if untreated before despatch, and as a concentrate if ore dressing operations are undertaken at or near the mine.

The commodity detail (involving eg the assayed metallic content of minerals) can be intricate and raises problems in achieving uniform definitional standards where various government agencies are involved. The field thus calls for a detailed knowledge of a wide variety of technical factors and a wide range of commodities as well as statistical expertise.

Administrative data are used, particularly as a primary source of lists of establishments. The principal source of data is generally direct censuses and surveys of establishments and (ideally) of the enterprises which own them.

Statistics of distributive trades and related services

key applications

Detailed information on the different types and sizes of establishments and their geographical location, is valuable for the urban and regional planners and for business organisations (eg for market research purposes, for planning location of new retail outlets, etc).

Thus, statistics of numbers engaged, value of sales and gross margins by kind of business (a classification of establishments according to the principal class of goods or services sold), size and location are desired, at least at intervals of 5 years or so.

Data on sales and gross margins form the basis of the *commodity flow approach* for estimating expenditures on gross domestic product and for compiling interindustry tables. Retail sales data, distributed according to kind of commodity, are used for weighting price series in indexes of retail prices in some countries. Others prefer a household expenditure survey for weighting purposes.

More frequent data on retail sales — annual, quarterly and monthly — are sensitive current economic indicators, providing forward indications of likely trends in manufacturing and other sectors of the economy.

Some countries also rely heavily on retail sales data classified by commodity or by kind of business for the preparation of estimates of the level and composition of personal consumption. Others endeavour to estimate personal consumption more directly, by way of household expenditure surveys.

what are distributive activities?

The basic notion seems to be to define the boundaries of this subject field in terms of the activity of reselling new and used goods without transforming them.

Industry classification boundary problems when businesses have mixed activities

This notion provides a criterion for drawing a line between the production and direct sale of goods by a manufacturer and the resale by wholesalers and retailers, who perform a service of distribution only.

But the line between the service of selling the goods and other kinds of services is not so easily drawn. Included are services, such as installation and repair, associated with the sale of goods, while more specialised related

activities such as transport, storage, financial, advertising and other business and personal services are excluded.

Characteristics of units in the distributive and service trades common to many, if not most, countries and which make statistical inquiries difficult, include the following:

- a **large number of small units;**
- the **rudimentary nature of records** kept by these units and, in some countries, the high proportion of illiterate respondents;
- the **high 'birth' and 'death' rate** as small businesses open and close and the frequent change of ownership and name of units — especially in retailing and service trades;
- **wide scatter of units** geographically;
- the **difficulty of identifying units** by observation — eg itinerant vendors, concessions granted by department stores etc.

One consequence of these characteristics, is the difficulty and high cost of maintaining a register which attempts to cover all units regardless of size. Another consequence is the need to employ sampling techniques, particularly area sampling, to capitalise on the fact that retail and many service trade units are likely to be highly correlated with the geographical distribution of the population.

where are the wholesalers?

Wholesalers are not easily identified by observation — especially agencies which maintain little or no stocks. Reconciling the various incomplete listing sources is generally a difficult task.

In general, distributive trade statistics are in the line of development of other industrial inquiries and should be closely integrated with them, both conceptually and operationally — eg by way of the use of a standard industrial classification and a common register (at least for larger units).

Statistics relating to the distributive trades and related services are generally among the least developed. Some countries have limited enquiries concentrating on the large stores, particularly in the cities, which account for a substantial and increasing share of wholesale and retail trade. The collection of a few essential data from the smaller units in the census and a sampling approach to the problem of collecting more elaborate or current statistics of distribution seems to be the way many countries proceed, once they have covered the larger units

Statistics of construction

construction activities create and repair unique products

The common attribute of the various activities listed as construction activities is the physical nature of their output. **Construction is essentially a goods producing industry** which has a good deal in common with manufacturing, **except that each product is unique** — at least in the sense of its construction on a different site — and it is this unique product characteristic that is the main source of difficulty for statistical operations in this field.

The construction industry's production usually involves the **creation of fixed assets or the repair and maintenance of fixed assets.**

Its product of capital equipment is of special interest also since its sales are an **alternative source for gross fixed capital expenditure estimates in respect of its clients.** It is usually the main source for capital expenditure on buildings.

Thus the industry is concerned with the creation or repair and maintenance of fixed assets on site, such as buildings, roads, railroads, aerodromes, irrigation projects, harbour or river works, water, gas, sewerage or stormwater drains or mains, electricity or other transmission lines, pipelines, oil refineries or other civil engineering projects.

boundary problems

As a whole it is a fairly well-defined field, although there are some mixed activity problems where enterprises engaged in mining, manufacturing, agriculture and transport, in particular, may also undertake substantial construction on their own account.

what is the construction 'establishment'?

Obviously the *establishment* definition (ie one activity, under one ownership at one location) will raise problems in the construction industry, since the activity of a construction firm takes place at a number of sites where successive jobs are undertaken, plus (perhaps) some off site workshop.

So the notion of location has to be understood in a special sense here. Generally it is necessary to define the construction establishment in terms of all the jobs of an enterprise in a particular statistical area, or all those handled from a particular base.

The UN recommendations are to use the **kind of activity unit,** which differs from the establishment in that it discards the single location restriction. Where possible this can be uniformly restricted to provide whatever regional subdivision of construction statistics is needed and feasible in each country so that, for example, large construction enterprises might be required to complete

census type returns in respect of all their construction activities in each region or local government area.

double counting of contractors' and sub-contractors' production?

Another major feature of the construction industry is **the prevalence of subcontracting.** The basic distinction is between the main contractor who undertakes the work for the investor and the subcontractors who work for the main contractors.

However, the situation in reality is fairly complex, as often the same enterprise will act as main contractor on one project but as a subcontractor on another.

Also, on a particular project, a main contractor may undertake all the work directly with its own labour or subcontract out the whole or part of the project. The subcontractors can thus be general builders and contractors or specialised enterprises who do only plumbing, electrical work, etc.

listing problems

Many small construction units go in and out of business in response to changing economic and seasonal factors. In addition they can change their address quite frequently and, moreover, often operate from premises, particularly dwellings, not readily identifiable as business offices. Identification of many of the statistical units involved in construction is therefore very difficult.

coverage approaches

In attempting to obtain a full record of the activities of the industry, the 'net' can be drawn from several points.

Thus it is possible to draw some information about the output of construction work from the industry's clients and again from government regulatory authorities. Various other sources may help to complete the closure — eg population censuses and labour force enquiries which provide information on labour inputs; taxation returns and enterprise surveys which might provide income and capital expenditure in respect of this industry (roughly distinguished on an enterprise unit basis from other industries covered).

A more satisfying but certainly more expensive approach is to collect all the information directly by census and surveys of the industry itself, supplemented by the other sources.

There is a need to coordinate the various partial enquiries in respect of coverage, concepts and definitions etc on an establishment or kind of activity units basis with (a) data on individual construction projects included in national development plans, and (b) macroeconomic aggregates for the construction sector as a whole.

Statistics of transport and other services

services — consumed as they are produced

In contrast with the construction industry, the output of these industries is services only — something generally consumed immediately it is produced. In general, transactions in services are more difficult to account for than transactions in goods and, in particular, accounting

and statistical systems do not deal adequately with flows of services between the establishments of enterprises.

TRANSPORTATION is a service produced by an industry and consumed by other industries and by final consumers. Statistics of transportation are thus statistics relating to the activities of establishments and enterprises providing the services.

what do we need to know about the transportation industry?

As with other production statistics, one is interested to know the structure of the industry, its contribution to product and the economic flows underlying this, such as the inputs of labour, capital and materials and its sales of services.

Similarly one is interested in the financial flows associated with the industry: funds for capital formation and outflows of income in the form of wages, dividends etc originating in transport enterprises.

Defined in these terms then, transport statistics relate to an industry and involve collecting data by censuses and surveys from the units which produce and/or use the services.

One would expect to obtain statistics on the same standard kinds of data items as for the other industries which have been discussed, plus whatever additional technical/engineering type data as may be peculiar to the transportation activity.

In this view of transportation, for the industries of land transport, water transport and air transport (subdivided further at the industry group level) we require details of commodities produced, ie different modes of transport subdivided by different categories of freight. As with other commodity statistics these can be recorded in money values or in physical quantity terms, such as passenger/kilometres or tonne/kilometres.

Thus we would put together here a picture of transportation as a subsystem of the economic system fitted into a general industrial classification and taking its appointed place within the framework of a system of national accounts. The industry can then be studied in the general context of economic theory at both macro and micro levels.

— about vehicle usage

An alternative to the employer registration-based census and survey approach is to use vehicle registration data as a frame for direct surveys of the owners and their vehicles. **Here the aim may be to provide the planners with basic data on traffic for estimating accident exposure rates and output in relation to equipment.**

what is the producing unit?

The principal problems which apply to the transport industry, and also to many other service industries, concern the definition of the statistical unit and its classification by principal economic activity (ie industry).

The first problem is that services are often delivered by mobile producing units (eg aircraft, ships and road and rail vehicles) which can only be located in terms of an area of operation (which may be international) or by way of a network (eg of depots and energy distribution or banking systems) **in which value added in the activity cannot be meaningfully attributed to individual locations of the enterprise.** In such industries, the only workable definition of the establishment may be the *kind of activity* (enterprise/industry) unit — ie all the enterprise's locations primarily engaged in that class of activity.

how to cover intra-enterprise services?

The second problem arises in the treatment of intra-enterprise services and in the definition of establishments and *ancillary units* in the case of services.

Thus the ISIC recognises the **flow of GOODS between units of an enterprise as a commodity flow** which is to be reported in industrial censuses and is to be **valued as if it were being sold by the one establishment of an enterprise to the other establishment.**

But the flow of SERVICES within an enterprise has generally not been treated in the same way in industry statistics (an exception is the case where one establishment provides the industrial type service of making up goods or the service of distributing goods, as in the case of a manufacturer's sales office holding stocks at a location separate from the factory).

Thus a potential establishment engaged in providing services within its own company or group only — eg carrying materials between its various plants and finished products to its sales offices and to its clients has been identified only as an ancillary unit to be treated as part of the industry of the parent producing unit(s) and having a separate identity only by virtue of its separate geographical location.

In effect, **the production of such services as transport, computing and the like have been recognised only as they are produced for SALE** as the primary product of an establishment classified to a service industry, or as the secondary product of an establishment classified to a goods industry or another service industry.

transport industry or transport activity statistics?

The implication of these conventions for the study of an activity like transportation (eg to analyse the demand for transportation equipment and roads) is that a good deal of value added etc that might be regarded as due to

transport and other service activity is being ascribed to goods producing or distributive trade industries.

Essentially the problem which the *ancillary unit* convention has tried to meet is that intra-enterprise flows of services are, generally speaking, not recorded in business accounts. Data on the activity are likely to be limited to employment, wages and capital equipment.

However, there is an increasing interest by businesses in accounting for these activities in market transaction terms and this is likely to be recognised in industry statistics and statistical unit standards. Thus the approach being introduced in Australia would lead to many such units being treated as establishments in their own right (coded to their own kind of activity) if the company maintains 'arms length' accounts for them. Other former ancillary units which do not meet the establishment 'availability of data' criteria will be treated as part of whatever establishment accounts for their activities (and remain in the same industry as they were as ancillary units).

This characteristic of industrial classification systems is of considerable significance in economic studies of the transportation activity since some of the larger enterprises maintain very large transport operations for their own use and almost every enterprise has some vehicles. Some countries in fact make special exceptions and treat the following *captive* units as transport establishments when they exceed a certain size —

(a) shipping lines

(b) airlines

(c) pipelines

(d) railway lines.

Conversely certain large repair and maintenance operations maintained by transport operators to repair their own equipment may be treated as establishments to be classified to appropriate manufacturing divisions, eg railway, ship and aircraft workshops.

Where the transport census is not carefully dovetailed with the other censuses of establishments after this fashion, problems will arise in making such classification transformations.

listing problems

Censuses and surveys may be the way to go for some of the service industries, but attempts to cover the transport industry comprehensively by way of censuses and surveys of road transport operators in particular, pose many problems of coverage since, in most countries, a substantial part of the industry is composed of small operators — one man with a truck or taxi, for example.

The coverage of **OTHER SERVICE INDUSTRIES**, eg legal, accounting, business services, personal services, is beset by problems of a similar kind to those encountered in transport statistics. International standards are very much in a formative stage, for lack of country experience in covering services systematically via production censuses and surveys.

financial services produced in non-geographical networks

While the **FINANCE INDUSTRY** is a service industry in ISIC, its contribution to GDP is not accounted for in the usual net output terms but is imputed from financial data on interest receipts and payments. In these terms, financial services are particularly difficult to relate to specific business locations. The producing unit in the case of a bank is really in the nature of a network so that there is little prospect or meaning in trying to determine value added by individual branches.

Sensibly, the establishment for the finance industries has to be defined as equivalent to the enterprise and data collections are undertaken in the very different context of enterprise level statistics.

Government services

government as a producer of goods and services

So far in identifying the boundaries of the fields of basic statistics underlying the production, consumption and capital formation accounts of the national accounts system, we have discussed industries producing goods and services, regardless of whether the establishments are owned by private businesses, governments or nonprofit institutions.

If we turn to ISIC, in order to see how government activities relate to the other goods producing and service industries that we have been studying, activities of government will be found under several industry divisions. The only division which is **exclusively** for government is division 91, *public administration and defence* which is now further subdivided in ISIC Rev. 3.

While many of the major parts of a central government, local government etc will be classified under this heading, other major public bodies may primarily engage in activities which should be classified to other categories of the international classification, eg social and related community services (such as education and health services), services incidental to transport, or agricultural services.

Of course the establishments of government business undertakings will be classified into such divisions as electricity, gas and water supply, transport, storage and communication and the like depending on the nature of their activity. **Apart from setting up the major division, public administration and defence, ISIC makes no**

distinction between establishments with regard to ownership.

There may be private businesses, nonprofit organisations and government undertakings in any of the other industrial categories.

It is not surprising that frequent misunderstandings arise when the size of government is discussed in the context of industry.

government as an institutional sector

The problem here is a confusion which arises from the fact that *government* is usually viewed as an institutional sector, such as the *general government sector*. In fact, few attempts have been made to identify establishment type units within government service agencies and to collect and publish appropriate data at that level.

'crossed wires' in public finance and industry statistics

Confusion between institutional and industry groupings has wider implications. Misunderstanding can occur between those who are endeavouring to organise statistical fields and collections on industry lines (eg mining statistics, manufacturing statistics, distributive and service trade statistics, construction statistics, government service statistics — as we have discussed economic statistics so far) and those who approach it on institutional lines (eg in terms of public authority statistics, company statistics, statistics of unincorporated businesses and households).

Those who conduct an industrial census may wish to include (say) the government munitions factory or government printing establishments in that system of statistics without regard for ownership while the public finance statisticians may bring them into their analysis of public accounts without regard to kind of activity. There is every likelihood here of 'crossed wires' in establishing definitional standards, classifications etc.

The SNA's dual sectoring arrangement, as depicted in Chapter 12, may be a useful device in sorting out the place of public finance statistics in the scheme of things. Thus there are really two related fields of government statistics; one which depicts the activity of government in producing goods and services, alongside other producers of goods and services, and the other which depicts the role of government as an institution raising funds to finance its service activities and making transfers to other sectors.

While accounts may be analysed for each public authority, each provincial government and for the central government, with perhaps a breakdown of central government activity by region, statistics of physical establishment-type units are generally not identified. Thus the subdivision of a government department seems likely to be limited for the present to what is available in the form of accounting units under which expenditure is classified by purpose.

In the long run the challenge to the public finance statistician is to persuade relevant government authorities to alter and improve the system of classification and other aspects of their accounting so that it provides a better economic policy instrument without conflicting with the needs of budgetary control (for example, by persuading them to use the standard economic and functional classification codes).

A last word on economic statistics

We have only skimmed across the great range and diversity of primary economic data and derived statistical systems that have been developed in different countries at different times. In doing so, we have tried to identify what it is that makes them 'systems' of statistics and where the possibilities lie for linking partial systems into more general purpose systems, given the common characteristics of money valuation which comprehensive systems of economic accounting try to exploit to the full.

In doing so, we have recognised that links between unified economic accounting systems and primary and special purpose derived statistical systems may be tenuous. Nevertheless there are obvious gains in analytical power and flexibility to be made from efforts to make the whole network of economic accounting and primary and derived data systems as coherent as possible.

Growing administrative and statistical capability combined with more complex information requirements of users will continue to move national statistical systems in this direction.

One of the keys to progress will be the constant pressure to resolve more and more comprehensive and detailed data confrontations of the kind we have discussed, and the responses to this by designers and managers of data processing systems responsible for programmes of both basic and derived statistics.

PART 4 : PRODUCTS — FOR THE SOCIAL ANALYST

CHAPTER 19
Social and demographic statistics — the elements of system

What is the relationship between economic and social statistics?
Establishing links with the national accounts
Attempts to define broad integrating frameworks — the SSDS
The social indicator movement
Focus on social concerns and population groups
Structural principles for frameworks
A final word on socio/demographic fields
Where to from here?
Appendix : 1 : Towards a System of Social and Demographic Statistics
2 : International Collection and Dissemination of Social Statistics in Statistical Publications

CHAPTER 20
Conceptual elements in social and demographic statistics

Statistical units and categories
Population stocks and flows
Classification by demographic and other attributes
Analysis of population characteristics
Households
Families
Married persons
Stocks and flows of families and households
The economically active and the inactive
Earning activities
Employed and unemployed
Participation rates and underemployment
Classifications by industry, occupation and industrial status
Special interest categories
Transactors/transactions
Social concerns and data items
Final thoughts on conceptual standards for social and demographic statistics

CHAPTER 21
Primary data systems for social and demographic statistics

Sources for social statistics
Administrative/statistical data bases — who is responsible?
Ultimate constraints
A last word on social statistics

CHAPTER 19

SOCIAL AND DEMOGRAPHIC STATISTICS — THE ELEMENTS OF SYSTEM

In Part 4 we move out of the economic accounting context to focus our attention on people ***as people*** *rather than as factors of production, owners, consumers and accumulators of wealth.*

We now find that the development of systems for social analysis becomes far more diverse once one goes beyond the domain of a money value accounting system and its various elaborations and extensions. It will never be simply a matter of filling more empty boxes within a progressively elaborated but essentially agreed framework.

In this chapter we discuss the search for system in social, labour and demographic fields and we find that this means working with a large number of separate systems — albeit loosely connected with economic accounting systems and their common money value denominator.

What is the relationship between 'economic' and 'social' statistics?

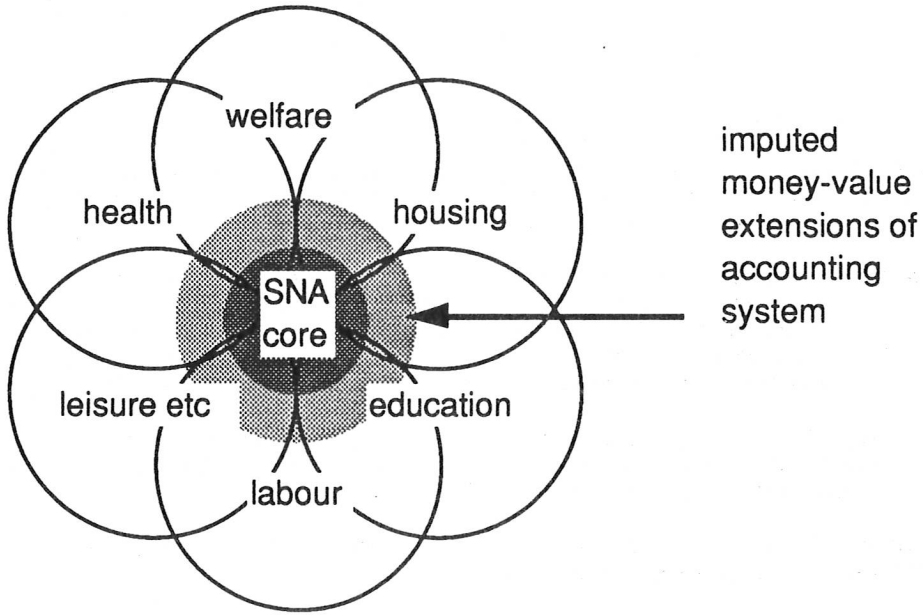
It is time to remind ourselves that this book is about national statistical systems and that in Parts 3 and 4 **we are looking for what is systematic about statistics of people living in society.**

In titling Parts 3 and 4 of our guided tour as, respectively, *Products — for the Economic Analyst* and *Products — for the Social Analyst* we may have given the impression that there is a simple dichotomy between economic analysis and social analysis.

Let us be clear at this point — there is no such dichotomy — economic analysis is simply a sub-set of social analysis and we should not try to divorce the two. However, it is legitimate to study statistics about society one step at a time. So having completed a guided tour of economic statistics we can now usefully concentrate our attention on the social statistics territory which lies outside the market, searching again for elements of systematic design.

The part of social statistics which is primarily concerned with data organising frameworks other than the SNA is generally labelled as *social statistics* or as *social and demographic statistics* — statistical fields which might be said to major on non-economic activities but which cannot ignore economic activities and data as they bear on the welfare of people.

In the diagram below, social and demographic statistics are represented by the circles indicated for representative fields of concern which overlap in various ways. The unshaded areas of these are the main focus of Part 4, since we toured the shaded areas in Part 3. The marginal territory of SNA extensions was visited in Chapter 13.



Establishing links with the national accounts

Accepting the fact that there is a limit to the usefulness of pursuing a national accounting perspective of the quality of life, social analysts seek systematic social, demographic and other information in alternative dimensions. At the same time linkages with the accounts are needed in order to facilitate cost/benefit and other financial policy considerations in the different fields of social concern.

Thus the money unit is bound to come into social and demographic statistics because of the ubiquitousness of cost issues governing the social services supplied to the population, whether these are provided by government or are purchased directly by the user.

One approach has been to carry over as much of the SNA standards and systems attributes as possible — eg to apply its concepts and classifications wherever appropriate as to statistical unit (eg persons, institutional sectors etc) and common operational arrangements (eg use of the same central register of statistical units) while recognising that, even if data cannot be brought into the comprehensive market valuation system via the money unit of measure, there may be other *numeraire*, such as person-years, numbers of incidents (eg arrests, consultations) and various physical units that will have additive properties up to a point.

One such strand of development aims to achieve some coherence and adding up properties in a social field by way of well-defined and consistent concepts and arithmetic identities after the manner of the SNA's relationship between income, product and expenditure and its summary of results in GDP and related aggregate measures of outcomes.

Statistics Canada's broad new conceptual approach to the organisation of health statistics illustrates this thinking. It aims to cover the spectrum from health outcomes to financial costs by building on a linkage of three main groups of concepts — life tables as generalized by the use of micro-simulation, I/O concepts especially related to techniques of production of health treatment, and SNA concepts of factor inputs, purposes of expenditure and sectors.

Its focus is to be on people and their healthfulness, with health outcomes as the major objective. Analogous to GDP or the CPI as overall indicators of the success of health programmes would be a 'population health expectancy' statistic with breakdowns across the population, for example by age, province, family composition, and income. The intention is that this should be capable of disaggregation, such that the dispersion or degree of inequality in health status across individuals can be indicated.

The other major focus, linking with the SNA, would be on (direct and indirect) economic costs — which should also be capable of disaggregation, eg by province, type of institution, purpose, financing arrangement, and type of resource used.

See Michael C Wolfson, A System of Health Statistics, Toward a New Conceptual Framework for Integrating Health Data, Review of Income and Wealth, March 1991

Thus the thinking behind such ambitious programmes as the proposed Canadian *System of Health Statistics* is that they should include the maintenance of a data base; the statistical data should be coherent and consistent; and the data base should be associated with analytical frameworks and mechanisms (eg to allow the derivation of the impacts of hypothetical scenarios). Clearly such frameworks will contain many *empty boxes* for a long time but may be influential in pointing the way for the progressive emergence and integration of appropriate data and theory.

Attempts to define broad integrating frameworks — The SSDS

Somewhat parallel thinking applies to various imaginative attempts which have been made by the United Nations and others to draw up some form of integrating framework within which the full range of social and demographic statistics might be developed into a cohesive system analogous to that for economic statistics.

While some useful approaches to this have been developed, none of them have achieved anything like the international acceptance of the SNA. It seems clear that once we go beyond the range of actual or imputed market valuation, with its common unit of account and its basis in near universal accounting conventions, we will necessarily move from a general system to a collection of more-or-less self-regarding data systems that can only be loosely connected.

Even so we do need to understand and strengthen the logical connections and to find better ways of invoking different statistical series in the solution of the inter-related problems that have constantly to be faced.

The value of doing this can be illustrated by considering Richard Stone's, *Towards a System of Social and Demographic Statistics* (SSDS) — United Nations 1975.

Stone set up social matrices in which opening states of the population , divided into groups by age and status (eg being in school, in the labour force or retired or in hospital), evolved through life sequences into a closing state. The life sequences and subsystems covered (see appendix 1) were —

(1) The size and structure of the population, births, deaths and migration

(2) Family formation, families and households

(3) Social class, stratification and mobility

(4) The distribution of income, consumption, accumulation and net worth

(5) Housing and its environment

(6) The allocation of time and the use of leisure

(7) Social security and welfare services

(8) Learning activities and education services

(9) Earning activities, employment services and the inactive

(10) Health and health services

(11) Public order and safety, offenders and their victims.

This is necessarily a multi-dimensional framework in which various indicators of the status and well-being of the population in different aspects of the social condition could be related in a larger framework of social, economic and demographic statistics and linked with the national accounts.

One of the dimensions is time (sub-sector 6 above). Time has long been used in statistics of hours of work. An analogy can be drawn between individuals' allocation of their time and the allocation of their monetary expenditure, and although 'time accounts' are not widely available, they are seen as a key to the development of more and better data on activity taking place beyond the SNA production boundary. Nevertheless units of time are not of equal value for different persons or for the same person at different points in the day or year. Consequently this is not a common denominator that can be aggregated with the same facility as money values.

The SSDS should not be regarded as a grand plan for assembling integrated social statistics within a single framework. In fact it has not been taken up by the international statistical community, possibly because of the very high costs of establishing its data base and because it did not satisfy the more sophisticated users (eg in that it was oriented solely to providing tabulations of data designed for use in particular life sequences).

This goes with this, goes with that...

Nevertheless, the SSDS has been a valuable aid to the statistical systems designer because of its objective and comprehensive explanation of logical connections between fields. But such considerations as government policy concerns and priorities and the division of portfolio responsibilities tend to be much more decisive in shaping the structure of a social statistics system.

The social indicator movement

Much of what follows in Part 4 is drawn from unpublished material, prepared by Mike Giles as a description of the full range of social statistics frameworks developed for Australia. This has since been published by ABS as A Guide to Australian Social Statistics, Catalogue No. 4160.0. This description of a representative country's system should be a valuable complement to the present 'guided tour', which can only offer pointers to the great diversity of approaches being followed by different countries.

In a swing away from the goal of comprehensive frameworks there has been a drive from the United Nations and the Organisation for Economic Cooperation and Development (OECD), in particular, to focus statistical development on producing aggregate comparative measures of individual well being, rather than to fill the empty boxes within such frameworks as the SSDS. This has been a useful movement in that it has directed attention at the *end results* or ultimate outputs of social betterment activities ie, their effects on well being.

One of the problems in this is that individual *well being* is an abstract notion and general agreement as to the value of different possible indicators of this is inevitably difficult.

For example, how was the the social statistician to compare the value attached by the community to a long and healthy life with other values such as freedom from want or public order and safety?

Nevertheless, the aspirations and aims of the social movement of the 60's and 70's needed to be expressed and the development of social indicators was a natural response.

goals, concerns and sub-concerns

The OECD defined 9 goal areas, within which concerns and sub-concerns were defined. For example, in goal area *Health*, a concern was **the probability of a healthy life through all stages of the life cycle** and a sub-concern within that concern was *length of life.*

In turn, the OECD defined social indicators which were to be direct and valid statistical measures to monitor **levels and changes over time in a fundamental social concern.**

The next step was to develop data collections from which to construct the indicators. But this was not successfully carried through — it was difficult to reach agreement on how to define the indicators in operational terms and it was asking too much of the community and their statistical services to produce the data which the theoretical concepts demanded.

As an interim measure, the OECD suggested a series of less demanding proxy indicators which at least utilized the available data and were a step along the way.

But there was also a progressive disenchantment with the social indicator approach as it became apparent that the indicators could not be sufficiently timely, disaggregated and program specific for government planning purposes.

Program planners therefore tended to ignore generalised social indicators and to demand statistics tailored to their specific programme needs. Thus their programmes were directed to particular target groups in the community — ie the frail aged, carers, the long term unemployed, single parents etc and they wanted these groups to be identified in the statistics. They needed specific information about them, such as age and birthplace, which were part of the eligibility criteria for many pensions and services, and they needed the information at regional as well as national levels.

About the same time the system approach of the SSDS was abandoned by the UN in favour of its FSDS — *Framework for Developing and Integrating Social and Demographic Statistics* (Studies in the Integration of Social Statistics: Technical Report, Series F No 24, United Nations 1979).

This was a list of tentative indicators under the SSDS topics, stopping well short of the original objectives of the social indicator movement, namely to appraise the qualitative non-economic aspects of social well being.

So the social indicator movement ran out of steam, at least in the impetus it gave to developing synoptic measures as objects of maximisation for each area of concern.

But the purpose behind the social indicators movement was not lost on social statisticians. Certainly, the deductive information requirements method adopted by the OECD was too ambitious

in the light of available statistical resources. Indicators based on broad goals and objectives simply could not in practice be disaggregated down to the program-specific levels of government planners and managers.

Nevertheless the point was made and today social need or purpose drives the development of social statistics much more firmly than before.

One lesson that has been learned is that a framework of social indicators is not a sufficient basis for the pursuit of matters of social concern, except as the indicators are related to well structured information systems about the circumstances and factors influencing them. More fully integrated frameworks such as the Statistics Canada System of Health Statistics, discussed earlier, give promise in this direction.

frameworks and systems

It might be noted that reference to a *framework* has come to imply something less ambitious than a *system* (which might have some, if not all, of the properties of the System of National Accounts, as indicated in Chapter 16). Some frameworks may be little more than a basis for juxtaposing a variety of information relevant to a particular social field — others will attempt to integrate the most common approaches to the organisation and presentation of statistics in the field.

An example of a relatively systematic framework is the FDES, a Framework for the Development of Environment Statistics, (United Nations Series M No. 78 1984). Thus, in a two way table, this relates the major components of the natural and man-made environment to a sequence of activities and events, environmental impacts resulting from these, and social responses to these impacts. This presentation indicates 'statistical topics' which are operational aspects of general environmental concern.

Focus on social concerns and population groups

The focus of data collection shifted from aggregate welfare measurement to the assessment of people-oriented *areas of social concern*, a term originally used by the OECD in its system of social indicators to measure *well being outputs*.

A response by many countries to the lack of common agreement on how concerns should be broken down into sub-concerns and as to what indicators should be selected as representative of these concerns has been simply to bring together all the information relevant to major areas of current social concern and to publish this in social reports on each topic (including, of course, such derived social indicators as can usefully be produced).

In more recent years there has been a focus also on the development of indicators of the social situation of *special population groups* such as women, children, youth, the elderly and disabled persons, with the same objective of assembling data in terms of current policy perspectives.

Structural principles for frameworks

At the same time, many countries have made a determined effort to establish frameworks specifically for each topic as a basis for their statistical development thrust.

For example, Australia saw particular value in an 'areas of social concern' framework which aligns reasonably well with the portfolio divisions of government. Thus, when it produces statistics for the social concerns 'health' or 'education' , they have a direct relationship to major government activity.

well-being status/ well-being instruments

The structuring effort within each concern area should not be confined to outputs but should also embrace social inputs such as government services and expenditure. Broadly speaking, the indicators will reflect, on the one hand, the aggregates of individual well-being (eg health and disability status) and on the other, community efforts to improve social conditions (eg health work force and expenditure by types of institution).

However, what may be required within one concern area may be significantly different from another. Thus, in the social concern *crime and justice* a structuring effort which recognises an offender, an offence and a victim and then traces the offender along the courts/prisons/parole path will be quite different from the structuring effort in the social concern areas of education or housing.

Again, countries will have different approaches in structuring their frameworks of statistics for social groups. To illustrate from Australian experience with the work on Youth, the structuring effort at its broadest level was the simple recognition of youth as a social group with a range of common problems and aspirations.

identifying the statistical units to be studied

The next structuring effort was to define the group (as persons aged between 15 and 24 years). This was the basic statistical unit in respect of which data items were to be counted.

and concerns to be analysed through data items

A further stage in establishing a conceptual framework was to decide how the problems and conditions of youth could be described. There were a number of choices here via areas of concern such as health, educational opportunities, opportunities for employment and so on.

and linkages with other frameworks

The process of putting issues into these types of compartments also raised the problem of links (ie education levels are linked to employment opportunities and this should be reflected in the framework).

At the level of *Youth health*, there was a further structuring of health into expectation of life/mortality, health status and health care. Thus, the framework for youth recognised the issue of health and examined health in terms of *dying, sickness and healthfulness* and *utilisation of health services by youth*. If data had been

available, a fourth category of *health resources* would also have been included.

common structures

The important point is that at every step there was a structuring effect at work. The Youth report was built on a framework. Not a single grand plan framework but one which was useful in throwing light on youth problems as such. In addition, elements of the framework for youth have been (and will be) used in other frameworks. In other words there needs to be deliberate elements of commonality in the various separate framework development thrusts.

the what and who dimensions

The two systems relating to *area of concern* and *social groups* have been connected in a number of ways. While statistics on (say) the disabled may encompass topics such as type of impairment, severity of handicap and type of care, some basic data about them have always been collected also on such general *concern areas* as income, education and housing.

There is thus a commonality of approach across each group in Australian social statistics. In the *Aged* publication for example, the social group *Aged* was considered from the point of view of health, employment, income and housing. Within the area of concern *health* the factors of mortality, health status (illness and handicap) and health care were considered.

Thus the framework underpinning the study of the Aged was similar to the framework used for Youth. Furthermore, the various classifications used within the common elements of the frameworks were *standard* classifications eg cause of death, labour force participation, decile classes of income.

To generalise, frameworks are a means to an end. They have to make a contribution to our statistical understanding of the real world. This may lead us into theoretical crudities and limitations but if the approach produces useful statistics as well as advancing the cause of standardisation, then it should be used.

In fact, there tends to be a constant process of integration of social statistics arising partly from the pressure to look at social change in a more holistic way and partly from the rapid advance in computer technology. Thus data from a number of collections can be arranged in an output data base to provide the user with 'all about' information on (say) retired and aged people. A data base on the aged might then include data on the disabled aged, family networks, housing arrangements, income and expenditure patterns, etc. The relatability of the data stems from using standard counting units and standard classifications, not from linkage of actual person records. We will pursue this matter of conceptual standards in Chapter 20.

Pressure for uniform standards in each field of social concern tends to come particularly from the UN Statistical Office, ILO, FAO, UNESCO and WHO and also from the Population Division of the UN Secretariat. The agencies

responsible for compiling internationally comparable statistics in each field are indicated in Appendix II.

A final word on socio/demographic fields

We have seen that once we pursue social concerns beyond the money-value accounting context we have much greater problems in making statistics of our social condition and behaviour *hang together.*

Precisely because we have no agreed measure of values for **beyond the production boundary** concerns we are probably bound to have to continue to work with a large number of separate systems — all loosely connected with the economic accounting systems which provide the only *across the board* communication connections and general perspectives.

The developing social statistics frameworks are the **frontier** areas of the national statistical system. They may be uncomfortable for systems developers but they hold the keys to the pertinence of an NSS to ultimate problems. The challenge facing the national statistical service is thus to identify issues of social concern and translate them into quantitative terms in a manner consistent with the society's value system.

Where to from here?

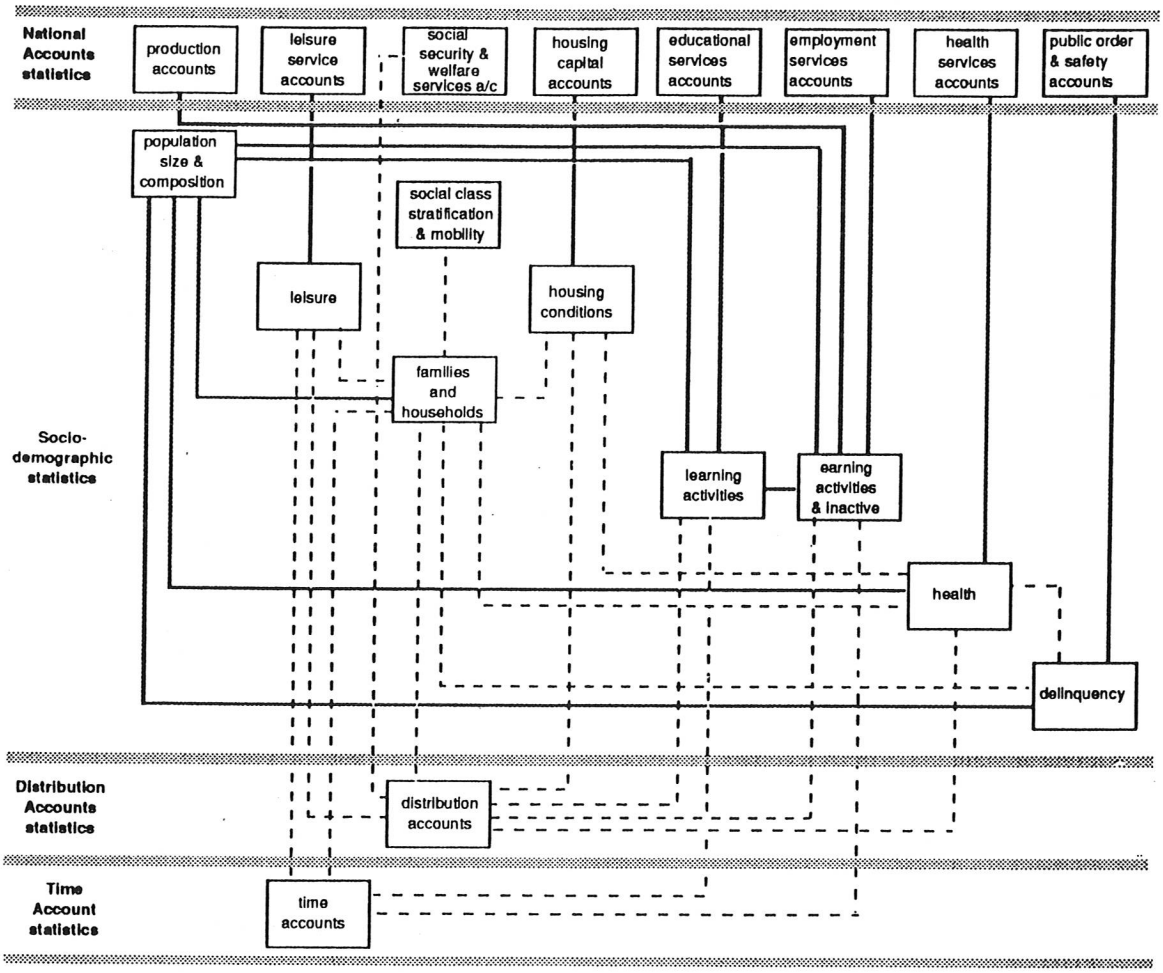
Against this background, we will look at the body of statistics available for social analysts from two viewpoints. In Chapter 20 we will look at the conceptual elements in social and demographic statistics — on the one hand, the basic statistical units to be identified and their classification into groupings of special interest; on the other hand, the social concerns to be explored and the basic data items relevant to these. These elements will be discussed from a user perspective, in terms of information requirements, conceptual standards and classifications and the identification of relevant social indicators.

Finally in Chapter 21, we will follow an 'information availability' approach and look at the nature of the major primary data collection systems supporting the social and demographic analyst, viz: a) Population census systems b) Household survey systems c) Administrative by-product systems d) Employer survey systems.

These will be discussed from a systems design and statistical operations perspective in order to identify the distinctive characteristics of the different kinds of data collection channels and the problems and possibilities of relating them in such contexts as discussed earlier.

APPENDIX I

(Towards a System of Social and Demographic Statistics, United Nations 1975, p21)



APPENDIX II

INTERNATIONAL COLLECTION AND DISSEMINATION OF SOCIAL STATISTICS IN STATISTICAL PUBLICATIONS

Fields of social concern	*Publication*	*Issuing organisation(b)*
	A. Specialised compilations	
Population, families and households	Demographic Yearbook	United Nations Secretariat — Statistical Office and Population Division
	Population and Vital Statistics reports	
	World population prospects	
	Ad hoc studies	
Housing and human settlements	Compendium of Human Settlements Statistics	United Nations Secretariat
Health, health services and nutrition(a)	World Health Statistics Annual, World Health Statistics Quarterly,	WHO
	Ad hoc studies	
	Food-balance sheets	FAO
	Reports on World Census of Agriculture	
	Review of Food Consumption Surveys	
	World Food Survey	
Learning and educational services(a)	Statistical Yearbook	UNESCO
	Ad hoc studies	
Earning activities a/	Yearbook of Labour Statistics	ILO
	Bulletin of Labour Statistics	
	Labour Force Estimates and Projections	
Income and consumption	*Ad hoc* studies	United Nations Secretariat
		ILO
		FAO
Social security and welfare services	Cost of Social Security	ILO
Public order and safety	Data collected quin-quinquenially and incorporated in reports for sessions of the Congress on Crime Prevention and Control,	United Nations Secretariat
	General statistics compendiums	Statistical Office

APPENDIX II — *continued*

INTERNATIONAL COLLECTION AND DISSEMINATION OF SOCIAL STATISTICS IN STATISTICAL PUBLICATIONS

Fields of social concern	*Publication*	*Issuing organisation(b)*
Leisure, culture and communications	Statistical Yearbook *Ad hoc* studies	UNESCO
Socio-economic groups and mobility	No international data collection	
	B. Multi-subject compilations	
	Statistical Yearbook Monthly Bulletin of Statistics World Statistics in Brief Compendium of Social Statistics Compendium of Statistics and Indicators on the Situation of Women *Ad Hoc* Studies	United Nations Secretariat Statistical Office
	World Tables, Vol. 11, Social Data World Development Report (Annex) Social Indicators of Development	World Bank
	Statistics on Children in UNICEF countries State of the World's Children (annex)	UNICEF
	Human Development Report (annex)	UNDP
	C. Economic statistics	
	National Accounts Statistics	United Nations Secretariat
	World Tables, Vol. 1, Economic Data	World Bank
	Government Finance Statistics	International Monetary Fund

(a) Statistics relevant to this topic from national demographic sources such as population censuses and vital statistics are also published in the Demographic Yearbook.

(b) In addition, comprehensive compilations of regional statistics, including social statistics, are issued by all of the regional commissions except ECE, in conjunction with their social and economic surveys. ECE issues a specialized compilation of statistics on housing, building and planning.

CHAPTER 20

CONCEPTUAL ELEMENTS IN SOCIAL AND DEMOGRAPHIC STATISTICS

We have seen in our guided tour of economic statistics, that the analytical value of economic accounting is vastly increased as it is disaggregated in terms of the groupings of statistical units (or transactors) in the economy and in terms of the data items (characteristics, activities, transactions) in respect of those units. In this chapter we pursue this notion as it applies more generally to statistics about society. We will look at some of the common conceptual elements in each of these two dimensions as they relate particularly to the circumstances of the people and population groups in society.

Statistical units and categories

The major statistical units in social statistics are persons, dwellings, households (derived from the relationship of persons in dwellings) and families.

In what follows, we will consider first the basic size and structure of the population in demographic terms, ie in terms of the units, characteristics and events that will govern changes in the number and disposition of people.

Then we will consider the further standard categorisations that relate to the activities of the people, to distinguish (eg) those who are economically active from those who are not.

Thirdly, we will look at further breakdowns of the population into categories which are of special interest for policy.

Population stocks and flows

In the language of statistics, a population can refer to a group of statistical units of any kind — dwellings, dogs, motor vehicles, whatever. In this discussion we are talking about populations of people, in the first instance in terms of individual persons, who of course might be differentiated into any number of categories to be counted — male, female, married, children, adults, patients, prisoners, students and so on— and in respect of whom we might collect any number of data items regarding their activities in social interaction.

Like other stocks and flows, population stocks relate to a point of time and population flows to a period of time. The opening stock of population at the beginning of a period may be linked with the closing stock at the end of the period if the additions to population by birth and in-migration and the deductions from the population by death and out-migration are known.

National statistical systems will be geared particularly to monitoring and forecasting population stocks and flows for the nation as a whole and for its major political divisions — for example, *population size* will be the key criterion in determining the levels of political representation and of financial grants for the provinces of a nation.

But there will be a need also for population stock and flow models of the **progression of people through the social system** as a whole or through any part of it.

Such models may identify populations in terms of standard demographic criteria, such as age and sex.

Depending on the aspect of life under study, they may also use special criteria to classify population into groups for which stocks and flows are needed. If we were interested in the flow of students through the educational system, for example, the specific criteria might be level and type of establishment attended, stage of work (for example, first and second stage of secondary school), subject of study and qualifications obtained at different stages of the education progression.

In developing appropriate population stock and flow models (such as the input-output and programming models illustrated in the SSDS) the main limitation is the availability of reliable data to fill the boxes of the stock /flow matrices required.

population stocks

The basic populations at a point of time for many such studies is established by taking a population census. Ten-yearly intervals for the census have been common in many countries in the past while some now have a five-yearly census. The United Nations speak of the *'1980 census'* and the *'1990 census'* embracing any census from 1975 to 1984 and any census from 1985 to 1994 respectively.

A census is designed as a simultaneous enumeration of population, based on a network which covers uniquely every habitation, and makes provision for covering the homeless and the migratory. Population may be based on the actual location of the population, (*de facto*, in which case every person is enumerated at the place where he was in the network at the census date), or based on the usual residence of the population. In some censuses, only those having legal permanent residence are counted, ie *de jure* residence. *Usual* residence involves counting persons intending to stay for more than a year (irrespective of legal status).

For some purposes, the population based on actual location may be wanted, eg in considering the requirements of a city which always has a large number of temporary visitors in its hotels.

The concept of *usual residence* may give a more useful population for many purposes, but it involves a working

definition of *usual residence*, which is increasingly difficult to apply in a mobile world.

flows in

To know the *additions to the population from births* requires a system of compulsory, universal registration of all live births occurring. Registration laws usually prescribe a network of registration areas and require that births be registered in the area of occurrence, which may be different from the area of residence.

Registration takes place after the event and total birth registrations in a period may differ from total births occurring in the same period. However, detailed tabulations of the statistics of births will enable re-allocation by place of residence of the parent and by date of occurrence, if such a degree of refinement is justified.

Additions by movements into the country as a whole may be known from counts at the frontier. They may also be classified by categories of persons arriving — for example, persons arriving to settle for the first time, and others returning after being away for at least one year. In respect of internal geographical divisions of the country, however, it would be rare to have statistics of all border crossings. There may be statistics of changes of residence which enable estimates of internal migration to be made.

flows out

The *reductions in population due to deaths* require a system of registration of all deaths, similar to that for births.

Losses due to migration again can be measured by statistics of total departures over the frontiers, or estimated from categories in these departures, for example by including only persons departing permanently, or for a known period exceeding a year.

Classification by demographic and other attributes

A population data base is much more than estimates of the population classified by geographical areas. The basic demographic structure of the population involves age and sex.

Classification by *sex* is usual in census tables. Classification by *age,* based on year of birth, or on age in completed years at the census, calls for a concentration into age groups, which, without being too numerous, will be narrow enough for particular purposes.

Geographic areas may also be classified, for example, into urban and rural, and by the size of the towns, or other population clusters.

There are other important attributes of the population such as *birthplace, national or ethnic origin*, which do not change with time, and others such as *marital status, fertility* (number of children born alive to women), *type of family* and *position in family, type of household* and *position in that household,* and *socio-economic class*, which may change with time.

For the statistical unit *family* one of the most important characteristics is *family type* (eg couple family with dependent children; with no dependent children; one parent family etc).

Updating census data by the technique applied to total population may also be carried out for the population by sex and by age, and could be performed for population by birthplace (which is invariable), given the requisite detail in the birth, death and migration statistics. Updatings by marital status and by families and households are conceivable, but present many difficulties.

Other socio-economic data gathered at a census enable the population to be classified into groups on the basis of characteristics of a non-demographic nature, such as by economic activity, educational activities and attainment, period of residence, labour force status, occupation, income, religion, health, family status and by type of housing.

Analysis of population characteristics

The geographical distribution of population is not only important as descriptive population statistics; it also provides a basis for comparison for many other statistics, provided the latter are available with the same geographical breakdown. Equally, the age structure statistics can be compared with other statistics which have an age dimension. The age structure itself is commonly represented diagrammatically by an *age pyramid.*

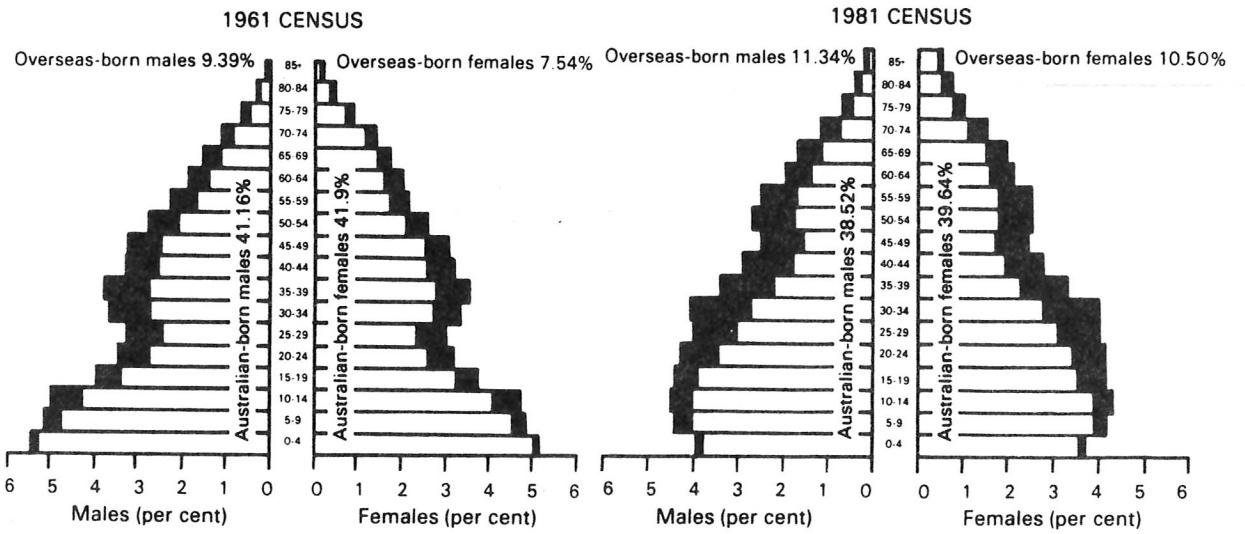
Age structures vary from place to place, especially between growth areas with many new arrivals and many young families, and long-established areas, with greater representation of older age groups. The numbers of women in the 15 to 44 age range has special significance for fertility. The numbers under 15 and over 65 are often grouped together as the 'dependent ages'; dependent on the 15 to 64 age group for their support.

The age pyramid also reflects the past history of the population, the effects of such happenings as waves of immigration, low birth rates during war or depression, and high mortality of males during war being reflected in the relative length of the steps of the pyramid.

elements in population projections

Present age structure is a necessary input to *projections of population* into the future. To make such projections, assumptions have to be made about future fertility, mortality and migration.

AGE PYRAMIDS OF AUSTRALIAN AND OVERSEAS BORN POPULATION, AUSTRALIA



Fertility is measured by relating number of births to mothers in an age group to the underlying population of all women in the same age group.

Mortality is measured by relating the number of deaths in each age group to the underlying total population in that age group, the two sexes being treated separately because of the persisting differences in their mortality rates, favouring females at most ages in many countries.

As measures of fertility, there is first the *crude birth rate*, simply the ratio of total births to total population; next the *general fertility* rate, relating total births to total women in the reproductive age; then there is *total fertility,* which is the sum of the age-specific fertilities over the reproductive age range, and represents the number of children born to a woman who lived through the reproductive years and experienced the current age-specific fertility rates at each age.

The *gross reproduction rate* is approximately half the *total fertility* and measures the *female* children born to a woman in the same circumstances; and *net reproduction rate* takes account of the fact that, at current female mortality rates, a proportion of girls and women will die before or during the reproductive age period.

To study overall current mortality independently of the current age structure, a *life table* is constructed from the current statistics of deaths and the underlying population at each age.

From the life table, the average expectation of life at birth of a generation of persons experiencing current age-specific mortality rates throughout their lives can be read; also the expectation of remaining life once any other age, for example, 10 years, 30 years or 60 years, has been reached.

Households

So far we have been considering population categories made up of individual persons with common characteristics of interest. But for many purposes of social and demographic analyses the more appropriate unit may be the household or the family.

The taking of a census itself groups most of the population into households, by the method of constructing a census network of areas, seeking out every dwelling in each area, and enumerating all persons either within, or usually resident, on census night.

what is a household?

Most censuses define a *household* as a *group of people living in the same dwelling.* As we noted in the previous chapter, it is usual to use the concept of *common housekeeping* arrangements in defining the household and sometimes common arrangements for eating are specified to help in the definition. Some censuses then define a *dwelling* as the living quarters of a household.

In a single structure, such as a house, there may be living arrangements for more than one household and a household may be a one person or a multi-person household. Thus a lodger who does not share the food of the main household is often defined as a separate household.

Sometimes, also, a person or persons having a *share of a house* are defined as a separate household. As can be imagined, there is a great variety of situations which the enumerator may encounter in a residential structure and the decision whether to recognise a separate household has to be made by the census enumerator on the spot, for which he or she must be carefully instructed.

Private households are groups of people who conduct a common housekeeping. There are other living arrangements which do not come within the definition of a private household. These are typically *institutional households* such as hospitals, where the housekeeping is in the hands of some kind of management.

Differences in concepts can produce differences in household grouping. If a census is taken on the basis of actual location, temporary visitors in the household become part of the household, while temporary absentees are omitted from the household.

Persons in institutional households, especially those of a short-term character such as hotels and short-stay hospitals will be more numerous at a census based on actual location than if the usual residence concept is used. The temporary displacements in a census based on actual location weaken the validity of household analysis.

Families

There are many people who live alone, making their own housekeeping arrangements, and the one-person household is a feature of any census. When two or more people live in one household then it is practicable to look for family relationships.

what constitutes a family?

The question whether a household contains a family will depend on the definition adopted for *family.*

There are two extremes of definition, one of which limits any one family to a couple and their unmarried children or to a lone parent and his or her unmarried children. This is the *nuclear* family.

At the other extreme a broader definition will include all persons with any blood or conjugal relationship living in the same household.

Of course, anyone may be a member of a family whose other members are not present. However, it is not practicable to try to gather all persons into families in this way at a census unless the absence is temporary.

In one household there may be enumerated *no family, one family* or *more than one family.* In the latter case it will be usual to select one family as the primary family, any other being a secondary family. Persons not in a family will be defined by their relationships in the household by such words as *boarder, visitor* or *partner.*

A narrow definition of family will result in larger numbers of families and more non-family members. A wider definition would reduce the numbers both of families and non-family members. Traditionally USA has used the wide definition of family at its census.

Sorting the members of a household into families and non-family members is done as part of the census processing on the evidence of family relationships given on the census schedules. One person will have been described as *head of the household* and the others will have indicated their relationship to that head.

In Australia, the concept of head of household has been dropped from the census and the first person listed on the census schedule becomes the 'reference' person to whom the relationship of other persons is indicated. Their relationship to each other is usually evident from the schedule, so that the census processors can identify families, although there are difficulties. The aim is usually to include de facto conjugal relationships between couples as family relationships and to include adopted children as in the family.

Married persons

By its question on marital status, the census establishes the numbers of married persons. This number is being added to by marriages and by the immigration of married persons and subtracted from by deaths, divorces and emigration. However, not all married couples live together and lone married persons enumerated at the census may be without their spouses permanently, or for an extended time, or only very temporarily in the case of a *de facto* census.

Stocks and flows of families and households

The flow data linking the stocks of families at successive times is not as straight-forward as the flow linking individuals, because families may be dissolved, by death, by divorce or by departure of a lone person's child. The survivors may then regroup with others, or not regroup, and where there was one family there may be none, or one, or two.

Households also can move from the *no-family* to the *one-family* status, or in reverse, as well as being dissolved, or established. Statistics needed to establish these stock and flow numbers are generally less than adequate. **Typically, the census or population survey will be called on to establish the stock. For flows the most important available statistics are those registrations of marriages, deaths and divorces and there is also an interest in surveying changes in the structure of families.**

The economically active and the inactive

At the point where interest focuses on the role of persons and households in the production process there is a need to distinguish categories in terms of their *involvement in economic activity.*

elements of labour statistics

Related concepts are *earning activities*, *employed* and *unemployed*, classifications of *industry*, *occupation* and *industrial status.*

Persons outside the labour force, the *economically inactive*, constitute a large proportion of the population. The largest category is that of dependent children. In countries where schooling is compulsory it is usual to consider labour force participation from the end of compulsory schooling, be it age 14 or 15 or 16. Below

that age, the population will be mostly *educationally active*, and accounted for in educational statistics. Preschool children may be considered separately and there may be statistics of the numbers who for various reasons live in institutions.

Of course, above the school leaving age full-time education continues for many. Other persons not economically active may be active in the household, or in leisure-type activities, or in voluntary work, and not seeking to enter the paid labour force.

Persons who have *independent means* have often been given a separate category, the intention being not to include them with persons living on retirement pensions.

There remain categories of persons who through illness or incapacity are permanently unable to work. Many of these may live in institutions, but others are in ordinary households. Further institutional populations such as long-term patients in hospitals and gaol inmates are temporarily withdrawn from the economically active population.

Separation by age is important in statistics of the inactive. From a labour force standpoint attention will be directed to the categories of the inactive who, although voluntarily out of the labour force, could possibly be available for first entry or re-entry at some future date. Interest also focuses on those discouraged job-seekers who are out of the labour force because of economic circumstances.

Earning activities

The concept of *earning activity* relates to any activity directed to the production of goods and services which is measurable in economic terms and in which, generally speaking, people are gainfully occupied.

In defining this, the 1982 *International Conference of Labour Statisticians* sought comparability with other economic and social statistics, including the SNA and its *boundary of production*. As discussed in Chapter 10 (pp 106 to 110) productive activity in these terms includes all production for the market and also production of goods for own consumption (as per the SNA's GDP concept), but not unpaid housework and *do-it-yourself* activity in households. It also includes the activities of government departments and private non-profit-making bodies which provide services which are not actually sold, but which can be measured in economic terms and classified into purpose categories such as education, medical care etc.

The *economically active population* (or *labour force*) thus *"comprises all persons who furnish the supply of labour for the production of economic goods and services as*

defined in the United Nations system of national accounts and balances during a specified time-reference period." Members of the labour force can be currently employed or unemployed.

Employed and unemployed

Thus the whole population can be divided into (1) *employed* or (2) *unemployed* or (3) *not in the labour force.* The employed include all those who had a job, even though they did not work at it. They can be subdivided into *persons in paid employment* and *self employed. Unemployed* are those with no job, available for work and actively seeking work.

Since the 1982 International Conference of Labour Statisticians an extended definition of the *unemployed* qualifies the restriction of *seeking work* as the governing criterion. It recognises the limited market for labour and the lack of labour exchanges and other facilities for actively seeking work in the rural areas of developing countries, in particular. This criterion is relaxed to include those currently available for employment, though not seeking it.

In Jamaica, for example, the unemployment rates in 1982 were 13% on the restricted definition, 27% on the extended definition.

The numbers of employed and unemployed are continually changing. Additions to the labour force come from young people leaving education, from immigrants and from women who were previously fully occupied in domestic duties. There are losses through retirement and death. The numbers of unemployed vary substantially with changes in the economic situation.

The member of the labour force who had a job in the reference period may be classified with reference to that job. For those who were unemployed, apart from those who are looking for their first job, it is common to classify them in accordance with the last job they had.

Participation rates and underemployment

Labour force participation rates are found by relating the employed and unemployed categories to the corresponding working age population.

To be useful, separate rates should be calculated for separate age-groups throughout the working ages. While differences in participation rates for male and females are becoming less and less marked, it is also useful to consider married women separately from single women. Substantially different patterns emerge, single women having higher participation rates and married women rates which vary more with age group than those for single women.

The employed labour force can also be usefully categorised according to (eg) the hours they were employed in a survey week or the months worked in the previous year so as to identify the *underemployed.* Thus underemployment exists when a person's employment is inadequate in relation to specified norms or alternative employment, account being taken of his or her occupational skill (training and working experience).

This was very significant in the Jamaican case cited above. The rate for unemployed (extended definition) plus underemployed was 43% of the labour force in 1982.

Classifications by industry, occupation and industrial status

The nature of the economic activity or job is described in terms of the industry in which it is performed and the occupation. The industrial establishment in which the job is held will be classified to an *industry* and the various kinds of work done by the labour force of the establishment will be classified as *occupations.*

A third classification will also be applied to the job, called *industrial status.* This classification separates out the person who owns a business and employs others, from the person who is an employee on wages or salary or some like method of remuneration. There are also categories for the self-employed person who has no employees, called *own account* and for the unpaid family worker who may be found in a small family business or farm.

The industrial classification used internationally is the *International Standard Industrial Classification of all Economic Activities* (ISIC), developed by the UN Statistical Commission. This is generally adopted as the basis for national classifications which may vary it in ways that will reflect the nature of the country's actual industrial structure but which will allow for re-compilation on an internationally comparable basis.

The occupational classification used internationally is the *International Standard Classification of Occupations* (ISCO), developed by the International Labour Organisation. Again there may be national versions with variations to meet national circumstances and needs.

industry relates to the kind of activity of a producing unit

occupation applies to tasks undertaken by workers

The *industry* is the branch of economic activity mainly performed by the establishment which supplies the job. *Occupation* tells something of the kind of work required of the person who fills it. The double classification by industry and occupation is necessary because many different kinds of work may occur in one establishment and most occupations can be carried on in many different industries.

The occupational classification may use degree of skill as one of the principles of its construction.

Occupation may also be a characteristic used in classifying persons by socio-economic groups. A detailed structural analysis of the labour force may show persons in occupational categories classified by age and by educational or trade qualifications. In some of the professional categories in the classification, responses to questions on qualifications may be used to assist in coding responses to the occupation questions.

For industries, the age of the labour force in each industry may be analysed, showing the industries which are absorbing the largest numbers of young workers.

Special interest categories

Just as the definition of the labour force and its division into the employed and unemployed is basic to the study of employment, so standard units definitions and classification categories are developed for each of the constituent fields of social and demographic statistics and become part of the specialised vocabulary of experts in each field.

As we have seen, one of the main paths to system in social statistics is through the identification of *population groups*. This can often be done by including the appropriate socio-economic variables in the data collection.

For example, women and men can be separately analysed via the variable *sex*, and older persons, adolescents and children by the variable *age*. Migrants need as a minimum the variables *birthplace* and *period of residence* in the country. Disadvantaged groups need *income* (as a minimum), together with *occupation* and *level of education*.

However, sufficient analytical power might not be available from the mere placement of a few explanatory variables in a range of surveys. Specific information peculiar to that population group is often needed.

For comparative purposes in the field of education statistics, for example, there is a need to establish unambiguous definitions of such population groups as eg.'new entrants' to the school system, 'repeaters', 'promotees' and 'drop outs'. Such categories may then need to be 'cross-classified' by basic fixed demographic characteristics such as age, sex and ethnic group.

Some purposes of analyses may require groups of the population to be differentiated on several special axes of interest — eg to assemble data about (say) single female parents of a particular level of education with a particular health history (eg of drug addiction). This may mean that cross-classification of persons by institution may be required — eg statistical analyses of former inmates of hospitals who received treatment for drug addiction might be the line of research for which confidentialised statistical aggregates might be needed.

Transactors/transactions

We have been discussing the categories of statistical units (*actors or transactors)* that interact in society and that need to be distinguished for useful analysis in pursuing different social concerns.

It will have become evident that once one goes beyond the basic unchanging age, sex, ethnic etc characteristics of individuals one is defining homogeneous categories in terms of the nature of their activities — *transactions or the outcome of transactions* of some kind.

In fact a great deal of social statistics consists simply of counts of the number of people in particular categories — we do not have a great deal of systematic data on the flows of transactions between transactor groups.

Nevertheless we still have a need to define and classify transactions — data items that relate to the activities/characteristics of our statistical units. Again there is the need to establish a common systematic vocabulary of concepts, definitions and classifications, so that a *'spade'* is always a spade whether in the minds of a survey respondent reading a questionnaire or in the minds of a user of statistics reading a publication or accessing an output data base.

Social concerns and data items

One such thrust will be a more or less direct extension of the SNA's data on the household sector.

> *Thus a natural response to a concern for groups which are disadvantaged in terms of income, levels of consumption and wealth would be to classify individuals (or households) according to these characteristics subdivided by eg. age and occupation (of the principal income earner).*

concepts of personal income and consumption

An important standard for such analyses might then be the *income class* levels to be used. Equal intervals will not necessarily be used over the whole range. Numbers in higher income groups tend to be small and therefore subject to large sampling error — it will be standard practice to combine them into broad groups for publication.

Concepts of *income* and *consumption* may also be modified to exclude elements that are difficult to cover by way of the household enquiries which enable breakdowns by detailed categories of individuals or households.

Thus it may be necessary to omit the imputations for income and expenditure attributed to the rental value of owner-occupied dwellings and the employers' contributions to pension funds. Again operational definitions would be necessary to discriminate between (1) the *total income* of various groups of the population and (2) *primary income* from participating in production

and current transfers and other benefits received, and (3) total *available household income* — total household income less direct taxes, social security and pension fund contributions.

Two concepts of household consumption are developed as social indicators — *outlays from total available household income* including in-kind consumption from own production, and *total consumption of the population* that is, personal consumption plus the value of goods and services furnished free of charge by governments, employers, and the like. The latter concept yields a more comprehensive measure of levels of living but may be impossible to apply to specific population categories.

There will obviously be other modifications and elaborations of SNA money-unit data item concepts and classifications in following through economic aspects of other fields of social concern — particularly to elaborate government and non-profit institution expenditure by purpose of expenditure. This will apply to the recommended standard social indicators and illustrative series as well as to the more *ad hoc* research requirements.

who sets the reporting definitions and classification standards?

The other likely standardisation thrust will relate to the great range of physical units that can apply to matters of social concern — measures of pollution, traffic flows, housing characteristics, energy stocks and flows and so on. The need to establish standards requires the response of scientists as well as government statisticians and the front-running is likely to be taken up as much by national standards agencies as by central statistical offices. Clearly the comprehensive systems interests of the latter will be a key input to the process.

While the development of standard social indicators within each of the traditional fields of social and demographic statistics has been a driving force in the establishment of both statistical unit and data item standards, some broad analytical thrusts, cutting across these, have also promoted the integration of standards.

Some major international strategies and events that have been influential in this way are the *United Nations Decade for Women*, the *International Year of Youth* (and the *Aged and Disabled Persons*), and *Health for All by the Year 2000*. Even broader interdisciplinary pressure has been generated by the United Nations *Third Development Decade* and the *Programme of Action for Agrarian Reform and Rural Development.*

Final thoughts on conceptual standards for social and demographic statistics

It is most important that, as we use bits of the information jig-saw in many different ways we are at least talking about the same thing. Certainly, an understanding of conceptual standards and vocabularies, such as are available in international standards manuals, is of great value in understanding and communicating.

In all this process of data standardisation for the development of national statistical information systems there tends to be a constant penetration of new areas of concern, resulting in the gradual firming up of standards in such a way that, ideally, the standards adopted for specific fields become a part of a more general system of standards. Necessarily the SNA standards will be very influential as the most general theoretical framework for associating information about people in social interaction. But there are many other dimensions to social interaction and there is a need for many other less comprehensive frameworks within which statistics can be contained and related.

Somehow there must be a continuous and lively process of ensuring that the technical terminology of statistical concepts and classifications and of data collection and programmes are in tune with the policy interests and vocabulary of users. The final chapter will consider the characteristics of the major data collections and programmes in this regard.

CHAPTER 21

PRIMARY DATA SYSTEMS FOR SOCIAL AND DEMOGRAPHIC STATISTICS

Few national statistical systems of broad scope are designed from the ground up to conform with such conceptual frameworks as were discussed in the previous two chapters.

Even with well-established statistical services the process is more often one in which the national statistician and/or the statisticians of major statistical units of government agencies are endeavouring to draw out and co-ordinate a variety of data from existing administrative sources with the help of a succession of censuses and surveys which plug gaps or provide integrating elements. The outcome is then something of an uneasy compromise between the ideals of standard conceptual systems, geared to an 'information requirements' approach to statistical operations, and the practicalities of accessing existing data sources, following an 'information available' approach. In this regard the social and economic statisticians have much in common! This chapter looks at some of the basic sources and channels of statistical data likely to be available for the kinds of purposes we have been discussing.

Sources for social statistics

Systems of primary statistics arise from the processing of a common set of source documents or records. We have noted the remarkable universality of the accounting records on which economic statistics are based and the facility which the common money-value denominator gives us in accumulating economic transactions data to whatever level of aggregation we want, from the individual transactor to the nation as a whole.

There will not be the same universal record keeping of events and outcomes — transactions, transformations, conditions etc — which are **not** measured in money value and so lie outside the conventions of accounting. Statistical aggregation will then be difficult:

(a) for lack of a common unit of measurement and

(b) for lack of comprehensive and consistent record keeping or of any strong natural incentive to seek out and conform to national standards.

Accounting data give rise to very current (eg monthly or weekly) statistics needed for monetary management. Non-economic social monitoring generally does not have this kind of immediacy even though the potential may be there for some series.

In general, non-economic social statistics will derive from:

- information which individuals can recall about themselves or hold in the form of certificates, approvals etc supplied by government agencies or which individuals are prepared to record for the government (eg in a diary form for a short period of time);

- the records kept by government and other institutions for their own specific administrative purposes (subject to government regulatory requirements).

lack of common standards is endemic in non-accounting sources

Except as statistical agencies can demonstrate the value of common standards or the need arises to share data with other administrative agencies, each agency is likely to introduce whatever definitions, classifications and methodologies they believe best suit their immediate administrative requirements or which have been decreed by their enabling legislation.

Some of the general attributes of each of the main sources of social and demographic data are considered below:

the big picture

1. POPULATION AND HOUSING CENSUSES

Most countries have made the decennial or quinquennial population and housing census a mandatory requirement on the national statistical authority. **From being essentially no more than counts of individuals and households classified by age, sex and ethnic origin, they have become our most comprehensive source of social and demographic data.**

The characteristic advantage of the census over other sources arises from its universal coverage, the wide range of data collected simultaneously, the well-tested and documented nature of the concepts and classifications used and the very wide range of possibilities for cross-classification by detailed geographical area and according to selected population characteristics.

The characteristic disadvantages lie in the long interval between data collections, and the logistic problems which constrain the complexity of the questionnaire and introduce delays, costs and difficulties in manipulating the records and producing timely detailed data.

Notwithstanding these disadvantages, the censuses are of paramount importance to any social and demographic monitoring programme.

They are a prime source for constructing key indicators, relating to social concerns and to special groups, and for the development of basic benchmark and reference data against which other social and demographic data can be compared and adjusted.

The strict limits on the amount and complexity of the information which it is expedient to try to collect on a census questionnaire have been circumvented to some extent in the census/survey programmes of some

countries by adding within-census or inter-censal sample collections to cover additional characteristics.

more flexible or current enquiries

2. HOUSEHOLD SAMPLE SURVEYS

To obtain social data from households more frequently and with respect to a greater range and complexity of information, many countries have developed systems of sample surveys that complement the population censuses and other sources.

Some of these are likely to be regular enquiries of a continuing representative sample of households with some topics, like labour force participation, being continuously monitored quarterly or monthly.

Additionally there are likely to be surveys of a *one time* or much less frequent nature dealing with particular issues in much greater depth.

Sometimes urban and rural surveys are taken separately — particularly for detailed enquiries of income and expenditure, or for comprehensive levels of living studies in countries where urban and rural living standards are very different.

Generally such sample surveys have the advantage of greater timeliness and flexibility, by comparison with periodical population and housing censuses.

Usually they will not be accurate enough to provide usable estimates for small data cells — they are likely to be designed only to give acceptable estimates for no more than an urban/rural split for major items. This may be quite acceptable for providing broad social indicators but limits the range of detailed analyses.

a continuing household survey system

While many surveys will continue to be undertaken on an ad hoc *one-off* basis to meet some special purpose need, there is a growing awareness of the value of **an integrated national household survey system.** This has been promoted particularly by the *United Nations Household Survey Capability Program* (UN HSCP). The characteristic approach and capability of such integrated survey systems is likely to be as follows:

— advantages

All surveys use the same base for their samples — usually derived from the population census. This allows scarce sampling expertise to be shared by all surveys. It also allows different surveys to be conducted in different blocks to minimise respondent fatigue and promote public cooperation.

Common geo-coding and other types of classification systems are also used to assist in relating data from the computer records of different surveys. Scarce software

and expertise for data editing and the compilation of results, including standard errors, can also be shared.

Uniform concepts and definitions of statistical units and data items are applied to assist in relating data from different surveys. Common coding frameworks can be used to save development resources and ensure compatibility.

Operational experiences and procedures can be shared and the best procedures adopted as standard.

A coherent long term programme of surveys can be developed, eg for 10 years, so that different projects are not competing for resources at the same time and that all necessary subject fields and topics are covered. In some cases this is done by adding different modules to a base survey (eg a labour force survey) in different years. In other cases, surveys are completely rotated in, for example, five year intervals.

A permanent field force can be maintained with a continuing workload. This allows them to build up expertise so that the amount of training for each survey is limited to new topics and procedural aspects.

institutional records

3. DIRECT SURVEYS OF INSTITUTIONAL POPULATIONS

Social conditions may also be monitored by statistical enquiries of institutions, such as schools or hospitals, rather than through households.

Sometimes this is done by a sampling of service centres (such as maternal and child health clinics, which in turn provide data relating to a sample of the mothers and children attending).

Such approaches may elicit very reliable data but in respect of a population defined by their contact with the institution — when perhaps it is the total population in each area which is the real object of interest.

registrations of events

4. CIVIL REGISTRATION AND REGISTERS

One of the major sources of social and demographic data is the legal requirement to register births, deaths, marriages and divorces.

The value of this channel is likely to depend very much on the *completeness* of the registration system (in many developing countries it is quite inadequate) and on the quality and accessibility of the records.

With computerisation in register offices, the extraction of reliable statistical data is becoming much less difficult. A characteristic difficulty — that events may be registered in geographical areas other than where the event

occurred — can be overcome by appropriate geographic coding and computer tabulation. The often poor quality of reporting or coding of information on the register (eg causes of death) calls for a continuing interest by national statisticians in the registration procedures and in the classification systems used.

other government records

5. OTHER GOVERNMENT ADMINISTRATIVE BY-PRODUCT DATA

A wide range of socio-economic data is collected by government agencies as part of their administrative and regulatory processes. The records of their relationship with people and other institutions, and the summary information prepared for monitoring and assessing these operations, can be a rich source of information for social and demographic analysts. **When such systems function well, both in an administrative and a statistical sense, they can generate timely data economically that are free from sampling error even when highly disaggregated. They may provide comparable time series stretching back over decades or even centuries.**

The challenge for the government statisticians is likely to be in persuading the agencies to apply common classifications, concepts, definitions and quality control procedures across the various fields, so that they can be used confidently in conjunction with one another.

Problems of timeliness, coverage and confidentiality are also likely to be encountered. Nevertheless, the objective should be to get the fullest possible value from the total stock of information in the records of government, within the constraints of costs and accepted guidelines on privacy and confidentiality.

private and semi-government sources

6. NON-GOVERNMENTAL ADMINISTRATIVE DATA

There are many private and semi government institutions with social data of a kind very similar to government agencies. Thus, private enterprises and non-private institutions, engaged in activities such as hospitals, schools, insurance and the like, may have a responsibility to maintain records and produce reports in accordance with government legislation providing for various forms of necessary government oversight of their activities.

It is important that the central statistical agency be involved in the drawing up or review of such arrangements in order that more general statistical requirements can be incorporated in the systems.

In countries where the coordinating role of the national statistician has not been well established, the flow of statistical information from such sources is likely to be unsystematic and incomplete. Sometimes the Statistician is forced to obtain such data by costly direct statistical surveys of non-government agencies, because the legislative requirements are out of line (perhaps quite arbitrarily) with the reporting standards which an integrated social reporting system would require.

Administrative/statistical data bases — who is responsible?

One of the tensions which arise in tapping statistical information from administrative sources is the extent to which the statistical agency should be involved in any necessary additional processing for statistical purposes.

Thus the strategy of the statistical office is likely to be that it should simply tap a point of concentration of existing information flows. However, if it is unable to ensure that the summarised data will meet its specifications it may have to agree to process the flow of unit records itself.

integrating administrative and statistical processes

A compromise may be for the central statistical agency to offer the services of its data processing systems and statistical standards experts to assist in designing a system which is an optimum one in terms of both a) the administrative agency's operating and management requirements and b) the general-purpose statistical requirements of the national statistical agency. **Generally it can be demonstrated that both the administrative agency and the statistical agency have a common interest in the two requirements being fully met.**

At the broadest level, governments may well insist that the central statistical authority and administrative agencies should jointly seek the optimum combination of administrative by-product channels and direct survey collections in respect of the individuals and organisations which are the object of interest.

The most efficient and effective system may well require the use of several optional channels in combination — eg to cover expenditure on motor vehicles for business and personal use by reference to i) purchasers, ii) sellers of motor vehicles, iii) motor vehicle registration authorities and iv) insurers.

In principle at least, the best arrangements for using administrative by-product data are likely to be ones in which only the administrative agency processes the administrative unit record but it does so to the coding, data entry and editing standards and summarisation specifications of the statistical agency.

Obviously close and continuing co-operation would be necessary to establish and maintain such a system but, over a period, the pay-off for all concerned is likely to be considerable.

Where the national statistical system is one in which there are separate institutions in each major subject matter field, each institution being responsible for both administrative data processing and official statistics in their field, the fullest and most economical exploitation of the data resources may result much more naturally. But, to the extent that social data systems inevitably extend across administrative agency boundaries, there will still be a major co-ordination role to be performed by

someone. Skills in negotiation may well be as important as skills in statistical systems design!

Ultimate constraints

At this point it is worth returning to the general issues raised in Chapter 2, relating to the political environment within which a national statistical system must operate.

In all the effort to gather and organise data about a society, whether directly by statistical survey or indirectly, through government administrative sources, the system must rest on a broad base of community support.

The community pays for the system in providing the necessary resources for the statistical service via the budget allocation and, increasingly, by purchasing the services. It pays for it also in the response it makes to requests for data which, in some degree, will be perceived as a burdon and an intrusion. **As attitudes change over time on such issues as 1) the right of government to intrude into the privacy of individuals and the affairs of business and of other centres of power and 2) the need for social controls and government intervention, it becomes a matter for political judgement and sensitive administration to ensure that an appropriate balance is maintained.**

A last word on social statistics

From this brief tour of the characteristics of the main primary sources of data about people in social interaction, two major conclusions seem to emerge —

(1) once one looks beyond the naturally comprehensive and systematic money-value accounting data usually available, the difficulties of developing national statistical systems evidently stem less from any lack of basic data than from their heterogeneity. Progress will depend very much on reducing this heterogeneity by developing and applying appropriate conceptual and operational standards.

(2) A basic condition for this is likely to be a widespread perception of the value of sharing data and off facilitating its systematic transformation into information about our society which will shed as much light as possible on the human condition and on the possibilities of improving it.

Obviously the state of development of national statistical systems will reflect the tension between the pursuit of self-interest by the individual and corporate actors in society and their acceptance of a social responsibility to provide statistical data. The production of social statistics is necessarily a social act in which the efforts of government statisticians, as intermediaries, is critical.

Recommended retail price $20.00

2011330012920
ISBN 0 642 17310 9